

**PATENT APPLICATION**

**COMPOSITIONS AND METHODS FOR DIAGNOSING AND  
TREATING MOOD DISORDERS**

Inventor(s): Huda Akil, a citizen of the United States, residing at  
3351 Windshadow Rd., Ann Arbor, MI 48105 USA

William E. Bunney, Jr., a citizen of the United States, residing at  
737 Kendall Drive, Laguna Hills, CA 92651 USA

Prabhakara V. Choudary, a citizen of India, residing at  
1522 Santa Rosa St., Davis, CA 95616 USA

Simon J. Evans, a citizen of the United States, residing at  
334 Ferman, Milan, Mi, 48160 USA

Edward G. Jones, a citizen of the United States, residing at  
P. O. Box 1259, Winters, CA 95694 USA

Jun Li, a citizen of the United States, residing at  
559 Loma Verde Ave., Palo Alto, CA 94304 USA

Juan F. Lopez, a citizen of the United States, residing at  
2680 Aspen Rd., Ann Arbor, MI 48108 USA

Robert C. Thompson, a citizen of the United States, residing at  
3208 Farmbrook Ct., Ann Arbor, MI 48108 USA

Richard Myers, a citizen of the United States, residing at  
957 Cottrell Way, Stanford, CA 94305 USA

Hiroaki Tomita, a citizen of Japan, residing at  
2019-A Los Trancos Dr., Irvine, CA, 92612

Marquis P. Vawter, a citizen of the United States, residing at  
28361 Via Alfonse Laguna, Niguel, CA 92677

Stanley Watson a citizen of the United States, residing at  
3351 Windshadow Rd., Ann Arbor, MI 48105

Entity: Small

TOWNSEND and TOWNSEND and CREW LLP  
Two Embarcadero Center, 8<sup>th</sup> Floor  
San Francisco, California 94111-3834  
Tel: 415-576-0200

## COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING MOOD DISORDERS

### CROSS-REFERENCES TO RELATED APPLICATIONS

5 [01] The present application claims the benefit of USSN 60/423,247, filed November 1, 2002 and USSN 60/431,454, filed December 6, 2002, the disclosures of which are hereby incorporated by reference in their entirety for all purposes.

### STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

10

[02] Not applicable.

### BACKGROUND OF THE INVENTION

[03] Clinical depression, including both bipolar disorders and major depression  
15 disorders, is a major public health problem, affecting an estimated 9.5% of the adult population of the United States each year. While it has been hypothesized that mental disorders, including mood disorders such as major depression and bipolar disorder as well as psychotic disorders such as schizophrenia, have complex genetic roots, little progress has been made in identifying gene sequences and gene products that play a role in causing these  
20 disorders, as is true for many diseases with a complex genetic origin (*see, e.g.*, Burmeister, *Biol. Psychiatry* 45:522-532 (1999)). Relying on the discovery that certain genes expressed in particular brain pathways and regions are likely involved in the development of mental disorders, the present invention provides methods for diagnosis and treatment of mental disorders, as well as methods for identifying compounds effective in treating mental  
25 disorders.

### BRIEF SUMMARY OF THE INVENTION

[04] In order to further understand the neurobiology of mood disorders such as bipolar disorders (BP) and major depression disorders (MDD), the inventors of the present  
30 application have used DNA microarrays to study expression profiles of human post-mortem brains from patients diagnosed with BP or MDD. The work has focused on three brain regions: the anterior cingulate cortex (AnCg), the dorsolateral prefrontal cortex (DLPFC), and the cerebellum (CB).

[05] The present invention demonstrates, for the first time, differential expression of the 72 nucleic acids listed in Table 2, the 16 nucleic acids listed in Table 3, or the 967 nucleic acids listed in Table 4, in the brains of patients suffering from mood disorders, such as bipolar disorder and major depression disorder, in comparison with normal control  
 5 subjects. In addition, the present invention identifies biochemical pathways involved in mood disorders, where the proteins encoded by the nucleic acids listed in Table 2, 3, or 4 are components of the biochemical pathways (e.g., the bFGF signal transduction pathway, the GPCR and cAMP/PI/Rho pathways, the proteasome pathway, the oxidative phosphorylation pathway, Myelination, Cytochrome P450, or the GABA and glutamate pathways; see also  
 10 Figures 1-5, 10-13, and 15).

[06] Finally, genes that are differentially expressed in MDD or BP and by gender are useful in diagnosing mood disorders, as the prevalence of certain mood disorders shows a gender bias. Differential expression by brain region similarly is a useful diagnostic and therapeutic tool, as certain mood disorders primarily affect certain brain regions.

15 [07] This invention thus provides methods for determining whether a subject has or is predisposed for a mental disorder such as bipolar disorder or major depression disorder. The invention also provides methods of providing a prognosis and for monitoring disease progression and treatment. Furthermore, the present invention provides nucleic acid and protein targets for assays for drugs for the treatment of mental disorders such as bipolar  
 20 disorder and major depression disorder.

[08] In some embodiments, the methods comprise the steps of: (i) obtaining a biological sample from a subject; (ii) contacting the sample with a reagent that selectively associates with a polynucleotide or polypeptide encoded by a nucleic acid that hybridizes under stringent conditions to a nucleotide sequence listed in Table 2, 3 or 4; and (iii)  
 25 detecting the level of reagent that selectively associates with the sample, thereby determining whether the subject has or is predisposed for a mental disorder.

[09] In some embodiments, the reagent is an antibody. In some embodiments, the reagent is a nucleic acid. In some embodiments, the reagent associates with a polynucleotide. In some embodiments, the reagent associates with a polypeptide. In some embodiments, the  
 30 polynucleotide comprises a nucleotide sequence of a gene listed in Table 2, 3, or 4. In some embodiment, the polypeptide comprises an amino acid sequence of a gene listed in Table 2, 3, or 4. In some embodiments, the level of reagent that associates with the sample is different (i.e., higher or lower) from a level associated with humans without a mental disorder. In

some embodiments, the biological sample is obtained from amniotic fluid. In some embodiments, the mental disorder is a mood disorder. In some embodiments, the mood disorder is selected from the group consisting of bipolar disorder and major depression disorder.

5 [10] The invention also provides methods of identifying a compound for treatment of a mental disorder. In some embodiments, the methods comprises the steps of: (i) contacting the compound with a polypeptide, which is encoded by a polynucleotide that hybridizes under stringent conditions to a nucleic acid comprising a nucleotide sequence of Table 2, 3, or 4; and (ii) determining the functional effect of the compound upon the  
10 polypeptide, thereby identifying a compound for treatment of a mental disorder.

[11] In some embodiments, the contacting step is performed *in vitro*. In some embodiment, the polypeptide comprises an amino acid sequence of a gene listed in Table 2, 3, or 4. In some embodiments, the polypeptide is expressed in a cell or biological sample, and the cell or biological sample is contacted with the compound. In some embodiments, the  
15 mental disorder is a mood disorder or psychotic disorder. In some embodiments, the mood disorder is selected from the group consisting of bipolar disorder I and II and major depression. In some embodiments, the psychotic disorder is schizophrenia. In some embodiments, the methods further comprise administering the compound to an animal, e.g., an animal subjected to stress as a model for depression and determining the effect on the  
20 animal, e.g., an invertebrate, a vertebrate, or a mammal. In some embodiments, the determining step comprises testing the animal's mental function.

[12] In some embodiments, the methods comprise the steps of (i) contacting the compound to a cell, the cell comprising a polynucleotide that hybridizes under stringent conditions to a nucleotide sequence of Table 2, 3, or 4; and (ii) selecting a compound that  
25 modulates expression of the polynucleotide, thereby identifying a compound for treatment of a mental disorder. In some embodiments, the polynucleotide comprises a nucleotide sequence listed in Table 2, 3, or 4. In some embodiment, the expression of the polynucleotide is enhanced. In some embodiments, the expression of the polynucleotide is decreased. In some embodiments, the methods further comprise administering the compound  
30 to an animal and determining the effect on the animal. In some embodiments, the determining step comprises testing the animal's mental function. In some embodiments, the mental disorder is a mood disorder or psychotic disorder. In some embodiments, the mood

disorder is selected from the group consisting of bipolar disorder I and II and major depression. In some embodiments, the psychotic disorder is schizophrenia.

[13] The invention also provides methods of treating a mental disorder in a subject. In some embodiments, the methods comprise the step of administering to the subject a therapeutically effective amount of a compound identified using the methods described above. In some embodiments, the mental disorder is a mood disorder or psychotic disorder. In some embodiments, the mood disorder is selected from the group consisting of bipolar disorder I and II and major depression. In some embodiments, the psychotic disorder is schizophrenia. In some embodiments, the compound is a small organic molecule, an antibody, an antisense molecule, aptamer, or a peptide.

[14] The invention also provides methods of treating mental disorders in a subject, comprising the step of administering to the subject a therapeutically effective amount of a polypeptide, which is encoded by a polypeptide that hybridizes under stringent conditions to a nucleic acid of Table 2, 3, or 4. In some embodiments, the polypeptide comprises an amino acid sequence encoded by a gene listed in Table 2, 3, or 4. In some embodiments, the mental disorder is a mood disorder or psychotic disorder. In some embodiments, the psychotic disorder is schizophrenia. In some embodiments, the mood disorder is a bipolar disorder or major depression.

[15] The invention also provides methods of treating mental disorders in a subject, comprising the step of administering to the subject a therapeutically effective amount of a polypeptide, wherein the polypeptide hybridizes under stringent conditions to a nucleic acid of Table 2, 3, or 4. In some embodiments, the mental disorder is a mood disorder or psychotic disorder. In some embodiments, the psychotic disorder is schizophrenia. In some embodiments, the mood disorder is a bipolar disorder or major depression.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[16] **Table 1:** Table 1 lists genes differentially expressed in mood disorder subjects.

[17] **Table 2:** Table 2 lists 72 genes differentially expressed in mood disorder subjects.

[18] **Table 3:** Table 3 lists 16 genes differentially expressed in specific brain regions and mood disorder.

[19] **Table 4:** Table 4 lists 967 genes differentially expressed in mood disorder subjects as determined by microarray analysis. Flag 1 indicates that the differential expression of the gene was confirmed by Real time PCR. Flag 2 indicates that differential expression of the gene was confirmed by anti-depressant studies. Flag 3 indicates that the gene belongs to an enriched gene ontology. Up and down indicates the direction of the changes compared to controls.

[20] **Table 5:** Table 5 lists Real time PCR results on sample genes that are differentially expressed in mood disorder subjects.

[21] **Table 6:** Table 6 lists anti-depressant treatment results for genes that are differentially expressed in mood disorder subjects.

[22] **Table 7:** Tables 7A-D lists the gene ontology of selected genes differentially expressed in mood disorder subjects.

[23] **Table 8:** Table 8 lists sample of genes that are differentially expressed in mood disorder subjects and are potential druggable targets.

[24] **Figure 1:** Figure 1 shows selected biochemical pathways for genes differentially expressed in mood disorder subjects.

[25] **Figure 2:** Figure 2 summarizes functions for signal transduction transcripts differentially expressed in MDD subjects.

[26] **Figure 3:** Figure 3 shows bFGF pathway transcripts differentially expressed in MDD subjects.

[27] **Figure 4:** Figure 4 shows values for differential expression of bFGF transcripts in MDD subjects.

[28] **Figure 5:** Figure 5 shows selected biochemical pathways that are dysregulated in mood disorders.

[29] **Figure 6:** Figure 6 shows selected biochemical pathways that are dysregulated in BP subjects.

[30] **Figure 7:** Figure 7 shows three genes overexpressed in mood disorder subjects that are located in the same chromosomal region.

[31] **Figure 8:** Figure 8 shows three genes overexpressed in mood disorder subjects that are located on 15q11-13 in the Prader-Willi region.

[32] **Figure 9:** Figure 9 shows certain genes regulated in human postmortem tissue and by antidepressants in rats.

[33] **Figure 10:** Figure 10 shows selected biochemical pathways (i.e., the GPCR and cAMP/PI/Rho pathways ) for genes differentially expressed in mood disorder subjects.

5 Two G protein coupled receptors, GPR37 and GPRC5B, are increased in both AnCg and DLPFC of BP patients, and decreased in MD. As downstream signaling pathways of GPCR, genes involved in cAMP pathway signaling are increased in BP patients, and decreased in MD. Genes involved in phosphatidylinositol pathways are deregulated specifically in MD.

[34] **Figure 11:** Figure 11 shows a selected biochemical pathway (i.e., the  
10 proteasome pathway) for genes differentially expressed in mood disorder subjects. The proteasome is an assembly of 28 alpha and beta subunits that functions to degrade proteins. The proteasome is involved in regulation of protein turnover and in particular oxidized proteins. There is an over representation of proteasome genes found in cortical regions of BP, but not in the cerebellum, suggesting that some functional compensation in the  
15 proteasome is occurring in BP patients.

[35] **Figure 12:** Figure 12 shows a selected biochemical pathway (i.e., the oxidative phosphorylation pathway) for genes differentially expressed in mood disorder subjects. The oxidative phosphorylation classification is involved in bioenergetics, metabolism, and as a byproduct can produce reactive oxygen species. This pathway is overly  
20 expressed in both bipolar and major depression, with differences between cortical regions and cerebellum.

[36] **Figure 13:** Figure 13 shows an example of a growth factor system (e.g., FGF) that is altered in mood disorders.

[37] **Figure 14:** Figure 14 shows RealTime PCR results which confirm that  
25 selected FGF-related genes first identified using microarray analysis are differentially expressed in mood disorders.

[38] **Figure 15:** Figure 15 shows selected genes in biochemical pathways involving GABA and glutamate that are differentially expressed in mood disorder subjects.

30

## DEFINITIONS

[39] A "mental disorder" or "mental illness" or "mental disease" or "psychiatric or neuropsychiatric disease or illness or disorder" refers to mood disorders (e.g., major

depression, mania, and bipolar disorders), psychotic disorders (e.g., schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder, brief psychotic disorder, and shared psychotic disorder), personality disorders, anxiety disorders (e.g., obsessive-compulsive disorder) as well as other mental disorders such as substance-related disorders, childhood disorders, dementia, autistic disorder, adjustment disorder, delirium, multi-infarct dementia, and Tourette's disorder as described in Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, (DSM IV). Typically, such disorders have a complex genetic and/or a biochemical component.

[40] "A psychotic disorder" refers to a condition that affects the mind, resulting in at least some loss of contact with reality. Symptoms of a psychotic disorder include, e.g., hallucinations, changed behavior that is not based on reality, delusions and the like. *See, e.g.,* DSM IV. Schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder, brief psychotic disorder, substance-induced psychotic disorder, and shared psychotic disorder are examples of psychotic disorders.

[41] "Schizophrenia" refers to a psychotic disorder involving a withdrawal from reality by an individual. Symptoms comprise for at least a part of a month two or more of the following symptoms: delusions (only one symptom is required if a delusion is bizarre, such as being abducted in a space ship from the sun); hallucinations (only one symptom is required if hallucinations are of at least two voices talking to one another or of a voice that keeps up a running commentary on the patient's thoughts or actions); disorganized speech (e.g., frequent derailment or incoherence); grossly disorganized or catatonic behavior; or negative symptoms, i.e., affective flattening, alogia, or avolition. Schizophrenia encompasses disorders such as, e.g., schizoaffective disorders. Diagnosis of schizophrenia is described in, e.g., DSM IV. Types of schizophrenia include, e.g., paranoid, disorganized, catatonic, undifferentiated, and residual.

[42] A "mood disorder" refers to disruption of feeling tone or emotional state experienced by an individual for an extensive period of time. Mood disorders include major depression disorder (i.e., unipolar disorder), mania, dysphoria, bipolar disorder, dysthymia, cyclothymia and many others. *See, e.g.,* Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, (DSM IV).

[43] "Major depression disorder," "major depressive disorder," or "unipolar disorder" refers to a mood disorder involving any of the following symptoms: persistent sad,



anxious, or “empty” mood; feelings of hopelessness or pessimism; feelings of guilt, worthlessness, or helplessness; loss of interest or pleasure in hobbies and activities that were once enjoyed, including sex; decreased energy, fatigue, being “slowed down”; difficulty concentrating, remembering, or making decisions; insomnia, early-morning awakening, or oversleeping; appetite and/or weight loss or overeating and weight gain; thoughts of death or suicide or suicide attempts; restlessness or irritability; or persistent physical symptoms that do not respond to treatment, such as headaches, digestive disorders, and chronic pain. Various subtypes of depression are described in, e.g., DSM IV.

[44] “Bipolar disorder” is a mood disorder characterized by alternating periods of extreme moods. A person with bipolar disorder experiences cycling of moods that usually swing from being overly elated or irritable (mania) to sad and hopeless (depression) and then back again, with periods of normal mood in between. Diagnosis of bipolar disorder is described in, e.g., DSM IV. Bipolar disorders include bipolar disorder I (mania with or without major depression) and bipolar disorder II (hypomania with major depression), see, e.g., DSM IV.

[45] An “agonist” refers to an agent that binds to a polypeptide or polynucleotide of the invention, stimulates, increases, activates, facilitates, enhances activation, sensitizes or up regulates the activity or expression of a polypeptide or polynucleotide of the invention.

[46] An “antagonist” refers to an agent that inhibits expression of a polypeptide or polynucleotide of the invention or binds to, partially or totally blocks stimulation, decreases, prevents, delays activation, inactivates, desensitizes, or down regulates the activity of a polypeptide or polynucleotide of the invention.

[47] “Inhibitors,” “activators,” and “modulators” of expression or of activity are used to refer to inhibitory, activating, or modulating molecules, respectively, identified using *in vitro* and *in vivo* assays for expression or activity, e.g., ligands, agonists, antagonists, and their homologs and mimetics. The term “modulator” includes inhibitors and activators. Inhibitors are agents that, e.g., inhibit expression of a polypeptide or polynucleotide of the invention or bind to, partially or totally block stimulation or enzymatic activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity of a polypeptide or polynucleotide of the invention, e.g., antagonists. Activators are agents that, e.g., induce or activate the expression of a polypeptide or polynucleotide of the invention or bind to, stimulate, increase, open, activate, facilitate, enhance activation or enzymatic

activity, sensitize or up regulate the activity of a polypeptide or polynucleotide of the invention, e.g., agonists. Modulators include naturally occurring and synthetic ligands, antagonists, agonists, small chemical molecules and the like. Assays to identify inhibitors and activators include, e.g., applying putative modulator compounds to cells, in the presence or absence of a polypeptide or polynucleotide of the invention and then determining the functional effects on a polypeptide or polynucleotide of the invention activity. Samples or assays comprising a polypeptide or polynucleotide of the invention that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of effect. Control samples (untreated with modulators) are assigned a relative activity value of 100%. Inhibition is achieved when the activity value of a polypeptide or polynucleotide of the invention relative to the control is about 80%, optionally 50% or 25-1%. Activation is achieved when the activity value of a polypeptide or polynucleotide of the invention relative to the control is 110%, optionally 150%, optionally 200-500%, or 1000-3000% higher.

[48] The term “test compound” or “drug candidate” or “modulator” or grammatical equivalents as used herein describes any molecule, either naturally occurring or synthetic, e.g., protein, oligopeptide (e.g., from about 5 to about 25 amino acids in length, preferably from about 10 to 20 or 12 to 18 amino acids in length, preferably 12, 15, or 18 amino acids in length), small organic molecule, polysaccharide, lipid, fatty acid, polynucleotide, oligonucleotide, etc. The test compound can be in the form of a library of test compounds, such as a combinatorial or randomized library that provides a sufficient range of diversity. Test compounds are optionally linked to a fusion partner, e.g., targeting compounds, rescue compounds, dimerization compounds, stabilizing compounds, addressable compounds, and other functional moieties. Conventionally, new chemical entities with useful properties are generated by identifying a test compound (called a “lead compound”) with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

[49] A “small organic molecule” refers to an organic molecule, either naturally occurring or synthetic, that has a molecular weight of more than about 50 Daltons and less than about 2500 Daltons, preferably less than about 2000 Daltons, preferably between about 100 to about 1000 Daltons, more preferably between about 200 to about 500 Daltons.

[50] "Determining the functional effect" refers to assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a polynucleotide or polypeptide of the invention (such as a polynucleotide of Table 2, 3, or 4 or a polypeptide encoded by a gene of Table 2, 3, or 4), e.g., measuring physical and chemical or phenotypic effects. Such functional effects can be measured by any means known to those skilled in the art, e.g., changes in spectroscopic (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein; measuring inducible markers or transcriptional activation of the protein; measuring binding activity or binding assays, e.g. binding to antibodies; measuring changes in ligand binding affinity; measurement of calcium influx; measurement of the accumulation of an enzymatic product of a polypeptide of the invention or depletion of an substrate; measurement of changes in protein levels of a polypeptide of the invention; measurement of RNA stability; G-protein binding; GPCR phosphorylation or dephosphorylation; signal transduction, e.g., receptor-ligand interactions, second messenger concentrations (e.g., cAMP, IP3, or intracellular  $\text{Ca}^{2+}$ ); identification of downstream or reporter gene expression (CAT, luciferase,  $\beta$ -gal, GFP and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, and ligand binding assays.

[51] Samples or assays comprising a nucleic acid or protein disclosed herein that are treated with a potential activator, inhibitor, or modulator are compared to control samples without the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation is achieved when the activity value relative to the control (untreated with activators) is 110%, more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

[52] "Biological sample" includes sections of tissues such as biopsy and autopsy samples, and frozen sections taken for histologic purposes. Such samples include blood, spinal fluid, sputum, tissue, lysed cells, brain biopsy, cultured cells, e.g., primary cultures, explants, and transformed cells, stool, urine, etc. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

[53] "Antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as the  
 5 myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

[54] An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair  
 10 having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain ( $V_L$ ) and variable heavy chain ( $V_H$ ) refer to these light and heavy chains respectively.

[55] Antibodies exist, e.g., as intact immunoglobulins or as a number of well-  
 15 characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce  $F(ab)'_2$ , a dimer of Fab which itself is a light chain joined to  $V_H$ - $C_H1$  by a disulfide bond. The  $F(ab)'_2$  may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the  $F(ab)'_2$  dimer into an Fab' monomer. The Fab' monomer is  
 20 essentially an Fab with part of the hinge region (see, Paul (Ed.) *Fundamental Immunology*, Third Edition, Raven Press, NY (1993)). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments  
 25 either produced by the modification of whole antibodies or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv).

[56] The terms "peptidomimetic" and "mimetic" refer to a synthetic chemical compound that has substantially the same structural and functional characteristics of the polynucleotides, polypeptides, antagonists or agonists of the invention. Peptide analogs are  
 30 commonly used in the pharmaceutical industry as non-peptide drugs with properties analogous to those of the template peptide. These types of non-peptide compound are termed "peptide mimetics" or "peptidomimetics" (Fauchere, *Adv. Drug Res.* 15:29 (1986); Veber and Freidinger *TINS* p. 392 (1985); and Evans *et al.*, *J. Med. Chem.* 30:1229 (1987), which are

incorporated herein by reference). Peptide mimetics that are structurally similar to therapeutically useful peptides may be used to produce an equivalent or enhanced therapeutic or prophylactic effect. Generally, peptidomimetics are structurally similar to a paradigm polypeptide (i.e., a polypeptide that has a biological or pharmacological activity), such as a CCX CKR, but have one or more peptide linkages optionally replaced by a linkage selected from the group consisting of, e.g.,  $-\text{CH}_2\text{NH}-$ ,  $-\text{CH}_2\text{S}-$ ,  $-\text{CH}_2\text{-CH}_2-$ ,  $-\text{CH=CH}-$  (cis and trans),  $-\text{COCH}_2-$ ,  $-\text{CH(OH)CH}_2-$ , and  $-\text{CH}_2\text{SO}-$ . The mimetic can be either entirely composed of synthetic, non-natural analogues of amino acids, or, is a chimeric molecule of partly natural peptide amino acids and partly non-natural analogs of amino acids. The mimetic can also incorporate any amount of natural amino acid conservative substitutions as long as such substitutions also do not substantially alter the mimetic's structure and/or activity. For example, a mimetic composition is within the scope of the invention if it is capable of carrying out the binding or enzymatic activities of a polypeptide or polynucleotide of the invention or inhibiting or increasing the enzymatic activity or expression of a polypeptide or polynucleotide of the invention.

[57] The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

[58] The term "isolated," when applied to a nucleic acid or protein, denotes that the nucleic acid or protein is essentially free of other cellular components with which it is associated in the natural state. It is preferably in a homogeneous state although it can be in either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is substantially purified. In particular, an isolated gene is separated from open reading frames that flank the gene and encode a protein other than the gene of interest. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure.

[59] The term "nucleic acid" or "polynucleotide" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of

natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions), alleles, orthologs, SNPs (haplotypes), and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); and Cassol *et al.* (1992); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

[60] The terms “polypeptide,” “peptide,” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymers. As used herein, the terms encompass amino acid chains of any length, including full-length proteins (i.e., antigens), wherein the amino acid residues are linked by covalent peptide bonds.

[61] The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline,  $\gamma$ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an  $\alpha$  carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. “Amino acid mimetics” refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[62] Amino acids may be referred to herein by either the commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical

Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[63] “Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, “conservatively modified variants” refers to those nucleic acids that encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations. Every nucleic acid sequence herein that encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid that encodes a polypeptide is implicit in each described sequence.

[64] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

[65] The following eight groups each contain amino acids that are conservative substitutions for one another:

- 1) Alanine (A), Glycine (G);
- 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);

- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
  - 7) Serine (S), Threonine (T); and
  - 8) Cysteine (C), Methionine (M)
- (see, e.g., Creighton, *Proteins* (1984)).

5 [66] "Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in  
10 both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

[67] The terms "identical" or percent "identity," in the context of two or more  
15 nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., 60% identity, optionally 65%, 70%, 75%, 80%, 85%, 90%, or 95% identity over a specified region), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence  
20 comparison algorithms or by manual alignment and visual inspection. Such sequences are then said to be "substantially identical." This definition also refers to the complement of a test sequence. Optionally, the identity exists over a region that is at least about 50 nucleotides in length, or more preferably over a region that is 100 to 500 or 1000 or more nucleotides in length.

25 [68] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence  
30 comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.



[69] A “comparison window”, as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith and Waterman (1970) *Adv. Appl. Math.* 2:482c, by the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity method of Pearson and Lipman (1988) *Proc. Nat'l. Acad. Sci. USA* 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel *et al.*, *Current Protocols in Molecular Biology* (1995 supplement)).

[70] An example of an algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1977) *Nuc. Acids Res.* 25:3389-3402, and Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the

sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

[71] The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

[72] An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequence.

[73] The phrase “selectively (or specifically) hybridizes to” refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent hybridization conditions when that sequence is present in a complex mixture (e.g., total cellular or library DNA or RNA).

[74] The phrase “stringent hybridization conditions” refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acid, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in

Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays"

(1993). Generally, stringent conditions are selected to be about 5-10° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength pH. The  $T_m$  is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at  $T_m$ , 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60° C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, optionally 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5X SSC, and 1% SDS, incubating at 42°C, or 5X SSC, 1% SDS, incubating at 65°C, with wash in 0.2X SSC, and 0.1% SDS at 65°C. Such washes can be performed for 5, 15, 30, 60, 120, or more minutes. Nucleic acids that hybridize to the genes listed in Tables 1-8 are encompassed by the invention.

[75] Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides that they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. Such washes can be performed for 5, 15, 30, 60, 120, or more minutes. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency.

[76] For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C

to about 65°C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis *et al.*, *PCR Protocols, A Guide to Methods and Applications* (1990).

[77] The phrase “a nucleic acid sequence encoding” refers to a nucleic acid that contains sequence information for a structural RNA such as rRNA, a tRNA, or the primary amino acid sequence of a specific protein or peptide, or a binding site for a trans-acting regulatory agent. This phrase specifically encompasses degenerate codons (i.e., different codons which encode a single amino acid) of the native sequence or sequences which may be introduced to conform with codon preference in a specific host cell.

[78] The term “recombinant” when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (nonrecombinant) form of the cell or express native genes that are otherwise abnormally expressed, under-expressed or not expressed at all.

[79] The term “heterologous” when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

[80] An “expression vector” is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

[81] The phrase “specifically (or selectively) binds to an antibody” or “specifically (or selectively) immunoreactive with”, when referring to a protein or peptide, refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated

5 immunoassay conditions, the specified antibodies bind to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised against a protein having an amino acid sequence encoded by any of the polynucleotides of the invention can be selected to obtain

10 antibodies specifically immunoreactive with that protein and not with other proteins, except for polymorphic variants. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays, Western blots, or immunohistochemistry are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. *See*, Harlow and Lane

15 *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, NY (1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity. Typically, a specific or selective reaction will be at least twice the background signal or noise and more typically more than 10 to 100 times background.

[82] One who is “predisposed for a mental disorder” as used herein means a person

20 who has an inclination or a higher likelihood of developing a mental disorder when compared to an average person in the general population.

## DETAILED DESCRIPTION OF THE INVENTION

### I. Introduction

[83] To understand the complex genetic basis of mental disorders, the present

25 invention provides studies that have been conducted to investigate the expression patterns of genes that are differentially expressed specifically in central nervous system of subjects with mood disorders. The large spectrum of symptoms associated with mental disorders is likely a reflection of the complex genetic basis and complex gene expression patterns in patients with

30 mental disorders. Different combinations of the genes disclosed herein can be responsible for one or more mental disorders. Furthermore, brain pathways or circuits as well as subcellular pathways are important for understanding the development and diagnosis of mental disorders. The selected brain regions described herein (AnCng, DLPFC, and CB) are implicated in the

clinical symptoms of mental disorders such as mood disorders. Brain imaging studies focusing on particular brain regions, cytoarchitectural changes in brain regions, expression of key neurotransmitters or related molecules in brain regions, and subcellular pathways in brain regions all contribute to the development of mental disorders, and thus are an important consideration in the diagnosis and therapeutic uses described herein.

[84] The present invention demonstrates the altered expression (either higher or lower expression) of the genes of Tables 1-8 at the mRNA level in selected brain regions of patients diagnosed with mood disorders (e.g., bipolar disorder and major depression disorder) in comparison with normal individuals. This invention thus provides methods for diagnosis of mental disorders such as mood disorders (e.g., bipolar disorder, major depression, and the like), psychotic disorders (e.g., schizophrenia, and the like), and other mental disorders by detecting the level of a transcript or translation product of the genes listed in Tables 1-8 as well as their corresponding biochemical pathways. The chromosomal location of such genes can be used to discover other genes in the region that are linked to development of a particular disorder.

[85] The invention further provides methods of identifying a compound useful for the treatment of such disorders by selecting compounds that modulates the functional effect of the translation products or the expression of the transcripts described herein. The invention also provides for methods of treating patients with such mental disorders, e.g., by administering the compounds of the invention or by gene therapy.

[86] The genes and the polypeptides that they encode, which are associated with mood disorders such as bipolar disease and major depression, are useful for facilitating the design and development of various molecular diagnostic tools such as GeneChips™ containing probe sets specific for all or selected mental disorders, including but not limited to mood disorders, and as an ante-and/or post-natal diagnostic tool for screening newborns in concert with genetic counseling. Other diagnostic applications include evaluation of disease susceptibility, prognosis, and monitoring of disease or treatment process, as well as providing individualized medicine via predictive drug profiling systems, e.g., by correlating specific genomic motifs with the clinical response of a patient to individual drugs. In addition, the present invention is useful for multiplex SNP or haplotype profiling, including but not limited to the identification of pharmacogenetic targets at the gene, mRNA, protein, and pathway level.

[87] The genes and the polypeptides that they encode, described herein, as also useful as drug targets for the development of therapeutic drugs for the treatment or prevention of mental disorders, including but not limited to mood disorders. Mental disorders have a high co-morbidity with other neurological disorders, such as Parkinson's disease or

5 Alzheimer's. Therefore, the present invention can be used for diagnosis and treatment of patients with multiple disease states that include a mental disorder such as a mood disorder.

[88] For example, antidepressants belong to different classes, e.g., desipramine, bupropion, and fluoxetine are in general equally effective for the treatment of clinical depression, but act by different mechanisms. The similar effectiveness of the drugs for  
10 treatment of mood disorders suggests that they act through a yet as unidentified common pathway. We disclose herein that different classes of antidepressants (specific serotonin reuptake inhibitors, like fluoxetine and tricyclic antidepressants, like desipramine) regulate a common gene, and/or a common group of genes as well as a unique set of genes when the human and animal results herein are compared.

## 15 II. General Recombinant nucleic acid methods for use with the invention

[89] In numerous embodiments of the present invention, polynucleotides of the invention will be isolated and cloned using recombinant methods. Such polynucleotides include, e.g., those listed in Tables 1-8, which can be used for, e.g., protein expression or during the generation of variants, derivatives, expression cassettes, to monitor gene  
20 expression, for the isolation or detection of sequences of the invention in different species, for diagnostic purposes in a patient, e.g., to detect mutations or to detect expression levels of nucleic acids or polypeptides of the invention. In some embodiments, the sequences of the invention are operably linked to a heterologous promoter. In one embodiment, the nucleic acids of the invention are from any mammal, including, in particular, e.g., a human, a mouse,  
25 a rat, a primate, etc.

### A. General Recombinant Nucleic Acids Methods

[90] This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (3rd ed. 2001); Kriegler, *Gene  
30 Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994)).

[91] For nucleic acids, sizes are given in either kilobases (kb) or base pairs (bp). These are estimates derived from agarose or acrylamide gel electrophoresis, from sequenced nucleic acids, or from published DNA sequences. For proteins, sizes are given in kilodaltons (kDa) or amino acid residue numbers. Proteins sizes are estimated from gel electrophoresis, from sequenced proteins, from derived amino acid sequences, or from published protein sequences.

[92] Oligonucleotides that are not commercially available can be chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage & Caruthers, *Tetrahedron Letts.* 22:1859-1862 (1981), using an automated synthesizer, as described in Van Devanter *et al.*, *Nucleic Acids Res.* 12:6159-6168 (1984). Purification of oligonucleotides is by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson & Reanier, *J. Chrom.* 255:137-149 (1983).

[93] The sequence of the cloned genes and synthetic oligonucleotides can be verified after cloning using, e.g., the chain termination method for sequencing double-stranded templates of Wallace *et al.*, *Gene* 16:21-26 (1981).

#### **B. Cloning Methods for the Isolation of Nucleotide Sequences Encoding Desired Proteins**

[94] In general, the nucleic acids encoding the subject proteins are cloned from DNA sequence libraries that are made to encode cDNA or genomic DNA. The particular sequences can be located by hybridizing with an oligonucleotide probe, the sequence of which can be derived from the sequences of the genes listed in Tables 1-8, which provide a reference for PCR primers and defines suitable regions for isolating specific probes. Alternatively, where the sequence is cloned into an expression library, the expressed recombinant protein can be detected immunologically with antisera or purified antibodies made against a polypeptide comprising an amino acid sequence encoded by a gene listed in Table 1-8.

[95] Methods for making and screening genomic and cDNA libraries are well known to those of skill in the art (*see, e.g.*, Gubler and Hoffman *Gene* 25:263-269 (1983); Benton and Davis *Science*, 196:180-182 (1977); and Sambrook, *supra*). Brain cells are an example of suitable cells to isolate RNA and cDNA sequences of the invention.

[96] Briefly, to make the cDNA library, one should choose a source that is rich in mRNA. The mRNA can then be made into cDNA, ligated into a recombinant vector, and



transfected into a recombinant host for propagation, screening and cloning. For a genomic library, the DNA is extracted from a suitable tissue and either mechanically sheared or enzymatically digested to yield fragments of preferably about 5-100 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in  
5 bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*, and the recombinant phages are analyzed by plaque hybridization. Colony hybridization is carried out as generally described in Grunstein *et al.*, *Proc. Natl. Acad. Sci. USA.*, 72:3961-3965 (1975).

[97] An alternative method combines the use of synthetic oligonucleotide primers  
10 with polymerase extension on an mRNA or DNA template. Suitable primers can be designed from specific sequences of the invention. This polymerase chain reaction (PCR) method amplifies the nucleic acids encoding the protein of interest directly from mRNA, cDNA, genomic libraries or cDNA libraries. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other *in vitro* amplification methods may also be  
15 useful, for example, to clone nucleic acids encoding specific proteins and express said proteins, to synthesize nucleic acids that will be used as probes for detecting the presence of mRNA encoding a polypeptide of the invention in physiological samples, for nucleic acid sequencing, or for other purposes (*see*, U.S. Patent Nos. 4,683,195 and 4,683,202). Genes amplified by a PCR reaction can be purified from agarose gels and cloned into an appropriate  
20 vector.

[98] Appropriate primers and probes for identifying polynucleotides of the invention from mammalian tissues can be derived from the sequences provided herein. For a general overview of PCR, *see*, Innis *et al. PCR Protocols: A Guide to Methods and Applications*, Academic Press, San Diego (1990).

25 [99] Synthetic oligonucleotides can be used to construct genes. This is done using a series of overlapping oligonucleotides, usually 40-120 bp in length, representing both the sense and anti-sense strands of the gene. These DNA fragments are then annealed, ligated and cloned.

[100] A gene encoding a polypeptide of the invention can be cloned using  
30 intermediate vectors before transformation into mammalian cells for expression. These intermediate vectors are typically prokaryote vectors or shuttle vectors. The proteins can be

expressed in either prokaryotes, using standard methods well known to those of skill in the art, or eukaryotes as described *infra*.

### III. Purification of proteins of the invention

[101] Either naturally occurring or recombinant polypeptides of the invention can be purified for use in functional assays. Naturally occurring polypeptides, e.g., polypeptides encoded by genes listed in Tables 1-8, can be purified, for example, from mouse or human tissue such as brain or any other source of an ortholog. Recombinant polypeptides can be purified from any suitable expression system.

[102] The polypeptides of the invention may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (*see, e.g.*, Scopes, *Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al.*, *supra*; and Sambrook *et al.*, *supra*).

[103] A number of procedures can be employed when recombinant polypeptides are purified. For example, proteins having established molecular adhesion properties can be reversibly fused to polypeptides of the invention. With the appropriate ligand, the polypeptides can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally the polypeptide can be purified using immunoaffinity columns.

#### A. Purification of Proteins from Recombinant Bacteria

[104] When recombinant proteins are expressed by the transformed bacteria in large amounts, typically after promoter induction, although expression can be constitutive, the proteins may form insoluble aggregates. There are several protocols that are suitable for purification of protein inclusion bodies. For example, purification of aggregate proteins (hereinafter referred to as inclusion bodies) typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells typically, but not limited to, by incubation in a buffer of about 100-150  $\mu$ g/ml lysozyme and 0.1% Nonidet P40, a non-ionic detergent. The cell suspension can be ground using a Polytron grinder (Brinkman Instruments, Westbury, NY). Alternatively, the cells can be sonicated on ice. Alternate methods of lysing bacteria are described in Ausubel *et al.* and Sambrook *et al.*, both *supra*, and will be apparent to those of skill in the art.

[105] The cell suspension is generally centrifuged and the pellet containing the inclusion bodies resuspended in buffer which does not dissolve but washes the inclusion bodies, e.g., 20 mM Tris-HCl (pH 7.2), 1 mM EDTA, 150 mM NaCl and 2% Triton-X 100, a non-ionic detergent. It may be necessary to repeat the wash step to remove as much cellular debris as possible. The remaining pellet of inclusion bodies may be resuspended in an appropriate buffer (e.g., 20 mM sodium phosphate, pH 6.8, 150 mM NaCl). Other appropriate buffers will be apparent to those of skill in the art.

[106] Following the washing step, the inclusion bodies are solubilized by the addition of a solvent that is both a strong hydrogen acceptor and a strong hydrogen donor (or a combination of solvents each having one of these properties). The proteins that formed the inclusion bodies may then be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to, urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents that are capable of solubilizing aggregate-forming proteins, such as SDS (sodium dodecyl sulfate) and 70% formic acid, are inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of the immunologically and/or biologically active protein of interest. After solubilization, the protein can be separated from other bacterial proteins by standard separation techniques.

[107] Alternatively, it is possible to purify proteins from bacteria periplasm. Where the protein is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to those of skill in the art (*see, Ausubel et al., supra*). To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM MgSO<sub>4</sub> and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

## **B. Standard Protein Separation Techniques For Purifying Proteins**

### **1. Solubility Fractionation**

[108] Often as an initial step, and if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol is to add saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This will precipitate the most hydrophobic proteins. The precipitate is discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, through either dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

### **2. Size Differential Filtration**

[109] Based on a calculated molecular weight, a protein of greater and lesser size can be isolated using ultrafiltration through membranes of different pore sizes (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

### **3. Column Chromatography**

[110] The proteins of interest can also be separated from other proteins on the basis of their size, net surface charge, hydrophobicity and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art.

[111] It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

#### IV. Detection of gene expression

5 [112] Those of skill in the art will recognize that detection of expression of polynucleotides of the invention has many uses. For example, as discussed herein, detection of the level of polypeptides or polynucleotides of the invention in a patient is useful for diagnosing mental disorders including mood disorders or psychotic disorders or a predisposition for a mood disorder or psychotic disorder. Moreover, detection of gene  
10 expression is useful to identify modulators of expression of the polypeptides or polynucleotides of the invention.

[113] A variety of methods of specific DNA and RNA measurement using nucleic acid hybridization techniques are known to those of skill in the art (*see*, Sambrook, *supra*). Some methods involve an electrophoretic separation (e.g., Southern blot for detecting DNA,  
15 and Northern blot for detecting RNA), but measurement of DNA and RNA can also be carried out in the absence of electrophoretic separation (e.g., by dot blot). Southern blot of genomic DNA (e.g., from a human) can be used for screening for restriction fragment length polymorphism (RFLP) to detect the presence of a genetic disorder affecting a polypeptide of the invention.

20 [114] The selection of a nucleic acid hybridization format is not critical. A variety of nucleic acid hybridization formats are known to those skilled in the art. For example, common formats include sandwich assays and competition or displacement assays. Hybridization techniques are generally described in Hames and Higgins *Nucleic Acid Hybridization, A Practical Approach*, IRL Press (1985); Gall and Pardue, *Proc. Natl. Acad. Sci. U.S.A.*, 63:378-383 (1969); and John *et al. Nature*, 223:582-587 (1969).  
25

[115] Detection of a hybridization complex may require the binding of a signal-generating complex to a duplex of target and probe polynucleotides or nucleic acids. Typically, such binding occurs through ligand and anti-ligand interactions as between a ligand-conjugated probe and an anti-ligand conjugated with a signal. The binding of the  
30 signal generation complex is also readily amenable to accelerations by exposure to ultrasonic energy.

[116] The label may also allow indirect detection of the hybridization complex. For example, where the label is a hapten or antigen, the sample can be detected by using antibodies. In these systems, a signal is generated by attaching fluorescent or enzyme molecules to the antibodies or in some cases, by attachment to a radioactive label (*see, e.g.,*

5 Tijssen, "*Practice and Theory of Enzyme Immunoassays*," *Laboratory Techniques in Biochemistry and Molecular Biology*, Burdon and van Knippenberg Eds., Elsevier (1985), pp. 9-20).

[117] The probes are typically labeled either directly, as with isotopes, chromophores, lumiphores, chromogens, or indirectly, such as with biotin, to which a

10 streptavidin complex may later bind. Thus, the detectable labels used in the assays of the present invention can be primary labels (where the label comprises an element that is detected directly or that produces a directly detectable element) or secondary labels (where the detected label binds to a primary label, e.g., as is common in immunological labeling). Typically, labeled signal nucleic acids are used to detect hybridization. Complementary

15 nucleic acids or signal nucleic acids may be labeled by any one of several methods typically used to detect the presence of hybridized polynucleotides. The most common method of detection is the use of autoradiography with  $^3\text{H}$ ,  $^{125}\text{I}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ , or  $^{32}\text{P}$ -labeled probes or the like.

[118] Other labels include, e.g., ligands that bind to labeled antibodies, fluorophores,

20 chemiluminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labeled ligand. An introduction to labels, labeling procedures and detection of labels is found in Polak and Van Noorden *Introduction to Immunocytochemistry*, 2nd ed., Springer Verlag, NY (1997); and in Haugland *Handbook of Fluorescent Probes and Research Chemicals*, a combined handbook and catalogue Published by Molecular Probes,

25 Inc. (1996).

[119] In general, a detector which monitors a particular probe or probe combination is used to detect the detection reagent label. Typical detectors include spectrophotometers, phototubes and photodiodes, microscopes, scintillation counters, cameras, film and the like, as well as combinations thereof. Examples of suitable detectors are widely available from a

30 variety of commercial sources known to persons of skill in the art. Commonly, an optical image of a substrate comprising bound labeling moieties is digitized for subsequent computer analysis.

[120] Most typically, the amount of RNA is measured by quantifying the amount of label fixed to the solid support by binding of the detection reagent. Typically, the presence of a modulator during incubation will increase or decrease the amount of label fixed to the solid support relative to a control incubation which does not comprise the modulator, or as compared to a baseline established for a particular reaction type. Means of detecting and quantifying labels are well known to those of skill in the art.

[121] In preferred embodiments, the target nucleic acid or the probe is immobilized on a solid support. Solid supports suitable for use in the assays of the invention are known to those of skill in the art. As used herein, a solid support is a matrix of material in a substantially fixed arrangement.

[122] A variety of automated solid-phase assay techniques are also appropriate. For instance, very large scale immobilized polymer arrays (VLSIPS™), available from Affymetrix, Inc. (Santa Clara, CA) can be used to detect changes in expression levels of a plurality of genes involved in the same regulatory pathways simultaneously. See, Tijssen, *supra.*, Fodor *et al.* (1991) *Science*, 251: 767- 777; Sheldon *et al.* (1993) *Clinical Chemistry* 39(4): 718-719, and Kozal *et al.* (1996) *Nature Medicine* 2(7): 753-759.

[123] Detection can be accomplished, for example, by using a labeled detection moiety that binds specifically to duplex nucleic acids (e.g., an antibody that is specific for RNA-DNA duplexes). One preferred example uses an antibody that recognizes DNA-RNA heteroduplexes in which the antibody is linked to an enzyme (typically by recombinant or covalent chemical bonding). The antibody is detected when the enzyme reacts with its substrate, producing a detectable product. Coutlee *et al.* (1989) *Analytical Biochemistry* 181:153-162; Bogulavski (1986) *et al. J. Immunol. Methods* 89:123-130; Prooijen-Knegt (1982) *Exp. Cell Res.* 141:397-407; Rudkin (1976) *Nature* 265:472-473, Stollar (1970) *Proc. Nat'l Acad. Sci. USA* 65:993-1000; Ballard (1982) *Mol. Immunol.* 19:793-799; Pisetsky and Caster (1982) *Mol. Immunol.* 19:645-650; Viscidi *et al.* (1988) *J. Clin. Microbiol.* 41:199-209; and Kiney *et al.* (1989) *J. Clin. Microbiol.* 27:6-12 describe antibodies to RNA duplexes, including homo and heteroduplexes. Kits comprising antibodies specific for DNA:RNA hybrids are available, e.g., from Digene Diagnostics, Inc. (Beltsville, MD).

[124] In addition to available antibodies, one of skill in the art can easily make antibodies specific for nucleic acid duplexes using existing techniques, or modify those antibodies that are commercially or publicly available. In addition to the art referenced

above, general methods for producing polyclonal and monoclonal antibodies are known to those of skill in the art (*see, e.g.,* Paul (3rd ed.) *Fundamental Immunology* Raven Press, Ltd., NY (1993); Coligan *Current Protocols in Immunology* Wiley/Greene, NY (1991); Harlow and Lane *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY (1988); Stites *et al.* (eds.) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Goding *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY, (1986); and Kohler and Milstein *Nature* 256: 495-497 (1975)). Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors (*see, Huse et al. Science* 246:1275-1281 (1989); and Ward *et al. Nature* 341:544-546 (1989)). Specific monoclonal and polyclonal antibodies and antisera will usually bind with a  $K_D$  of at least about 0.1  $\mu\text{M}$ , preferably at least about 0.01  $\mu\text{M}$  or better, and most typically and preferably, 0.001  $\mu\text{M}$  or better.

[125] The nucleic acids used in this invention can be either positive or negative probes. Positive probes bind to their targets and the presence of duplex formation is evidence of the presence of the target. Negative probes fail to bind to the suspect target and the absence of duplex formation is evidence of the presence of the target. For example, the use of a wild type specific nucleic acid probe or PCR primers may serve as a negative probe in an assay sample where only the nucleotide sequence of interest is present.

[126] The sensitivity of the hybridization assays may be enhanced through use of a nucleic acid amplification system that multiplies the target nucleic acid being detected. Examples of such systems include the polymerase chain reaction (PCR) system, in particular RT-PCR or real time PCR, and the ligase chain reaction (LCR) system. Other methods recently described in the art are the nucleic acid sequence based amplification (NASBA, Cangene, Mississauga, Ontario) and Q Beta Replicase systems. These systems can be used to directly identify mutants where the PCR or LCR primers are designed to be extended or ligated only when a selected sequence is present. Alternatively, the selected sequences can be generally amplified using, for example, nonspecific PCR primers and the amplified target region later probed for a specific sequence indicative of a mutation.

[127] An alternative means for determining the level of expression of the nucleic acids of the present invention is *in situ* hybridization. *In situ* hybridization assays are well known and are generally described in Angerer *et al., Methods Enzymol.* 152:649-660 (1987). In an *in situ* hybridization assay, cells or tissue, preferentially human cells or tissue from the



cerebellum or the hippocampus, are fixed to a solid support, typically a glass slide. If DNA is to be probed, the cells are denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of specific probes that are labeled. The probes are preferably labeled with radioisotopes or fluorescent reporters.

## 5 V. Immunological detection of the polypeptides of the invention

[128] In addition to the detection of polynucleotide expression using nucleic acid hybridization technology, one can also use immunoassays to detect polypeptides of the invention. Immunoassays can be used to qualitatively or quantitatively analyze polypeptides. A general overview of the applicable technology can be found in Harlow & Lane, *Antibodies: A Laboratory Manual* (1988).

### A. Antibodies to target polypeptides or other immunogens

[129] Methods for producing polyclonal and monoclonal antibodies that react specifically with a protein of interest or other immunogen are known to those of skill in the art (*see, e.g.,* Coligan, *supra*; and Harlow and Lane, *supra*; Stites *et al., supra* and references  
15 cited therein; Goding, *supra*; and Kohler and Milstein *Nature*, 256:495-497 (1975)). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors (*see, Huse et al., supra*; and Ward *et al., supra*). For example, in order to produce antisera for use in an immunoassay, the protein of interest or an antigenic fragment thereof, is isolated as described herein. For example, a  
20 recombinant protein is produced in a transformed cell line. An inbred strain of mice or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen.

[130] Polyclonal sera are collected and titered against the immunogen in an  
25 immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Polyclonal antisera with a titer of  $10^4$  or greater are selected and tested for their cross-reactivity against unrelated proteins or even other homologous proteins from other organisms, using a competitive binding immunoassay. Specific monoclonal and polyclonal antibodies and antisera will usually bind with a  $K_D$  of at least about 0.1 mM, more usually at  
30 least about 1  $\mu$ M, preferably at least about 0.1  $\mu$ M or better, and most preferably, 0.01  $\mu$ M or better.

[131] A number of proteins of the invention comprising immunogens may be used to produce antibodies specifically or selectively reactive with the proteins of interest.

Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Naturally occurring protein, such as one comprising an amino acid sequence encoded by a gene listed in Table 1-8 may also be used either in pure or impure form. Synthetic peptides made using the protein sequences described herein may also be used as an immunogen for the production of antibodies to the protein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells and purified as generally described *supra*. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated for subsequent use in immunoassays to measure the protein.

[132] Methods of production of polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the polypeptide of interest. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (*see*, Harlow and Lane, *supra*).

[133] Monoclonal antibodies may be obtained using various techniques familiar to those of skill in the art. Typically, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see*, Kohler and Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include, e.g., transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse *et al.*, *supra*.

[134] Once target protein specific antibodies are available, the protein can be measured by a variety of immunoassay methods with qualitative and quantitative results

available to the clinician. For a review of immunological and immunoassay procedures in general *see*, Stites, *supra*. Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in Maggio *Enzyme Immunoassay*, CRC Press, Boca Raton, Florida (1980); Tijssen, *supra*; and Harlow and Lane, *supra*.

[135] Immunoassays to measure target proteins in a human sample may use a polyclonal antiserum that was raised to the protein (e.g., one has an amino acid sequence encoded by a gene listed in Table 1-8) or a fragment thereof. This antiserum is selected to have low cross-reactivity against different proteins and any such cross-reactivity is removed by immunoabsorption prior to use in the immunoassay.

#### **B. Immunological Binding Assays**

[136] In a preferred embodiment, a protein of interest is detected and/or quantified using any of a number of well-known immunological binding assays (*see, e.g.*, U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, *see also* Asai *Methods in Cell Biology Volume 37: Antibodies in Cell Biology*, Academic Press, Inc. NY (1993); Stites, *supra*. Immunological binding assays (or immunoassays) typically utilize a “capture agent” to specifically bind to and often immobilize the analyte (in this case a polypeptide of the present invention or antigenic subsequences thereof). The capture agent is a moiety that specifically binds to the analyte. In a preferred embodiment, the capture agent is an antibody that specifically binds, for example, a polypeptide of the invention. The antibody may be produced by any of a number of means well known to those of skill in the art and as described above.

[137] Immunoassays also often utilize a labeling agent to specifically bind to and label the binding complex formed by the capture agent and the analyte. The labeling agent may itself be one of the moieties comprising the antibody/analyte complex. Alternatively, the labeling agent may be a third moiety, such as another antibody, that specifically binds to the antibody/protein complex.

[138] In a preferred embodiment, the labeling agent is a second antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second antibody can be modified with a detectable moiety, such as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin.

[139] Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G, can also be used as the label agents. These proteins are normal constituents of the cell walls of streptococcal bacteria. They exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see*,  
 5 generally, Kronval, *et al. J. Immunol.*, 111:1401-1406 (1973); and Akerstrom, *et al. J. Immunol.*, 135:2589-2542 (1985)).

[140] Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, preferably from about 5 minutes to about 24 hours. The incubation time will depend  
 10 upon the assay format, analyte, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 10°C to 40°C.

### 1. Non-Competitive Assay Formats

[141] Immunoassays for detecting proteins of interest from tissue samples may be  
 15 either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of captured analyte (in this case the protein) is directly measured. In one preferred "sandwich" assay, for example, the capture agent (e.g., antibodies specific for a polypeptide encoded by a gene listed in Table 1-8) can be bound directly to a solid substrate where it is immobilized. These immobilized antibodies then capture the polypeptide present in the test  
 20 sample. The polypeptide thus immobilized is then bound by a labeling agent, such as a second antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second can be modified with a detectable moiety, such as biotin, to which a third labeled molecule can specifically bind, such as enzyme-  
 25 labeled streptavidin.

### 2. Competitive Assay Formats

[142] In competitive assays, the amount of analyte (such as a polypeptide encoded by a gene listed in Table 1-8) present in the sample is measured indirectly by measuring the amount of an added (exogenous) analyte displaced (or competed away) from a capture agent  
 30 (e.g., an antibody specific for the analyte) by the analyte present in the sample. In one competitive assay, a known amount of, in this case, the protein of interest is added to the sample and the sample is then contacted with a capture agent, in this case an antibody that

specifically binds to a polypeptide of the invention: The amount of immunogen bound to the antibody is inversely proportional to the concentration of immunogen present in the sample. In a particularly preferred embodiment, the antibody is immobilized on a solid substrate. For example, the amount of the polypeptide bound to the antibody may be determined either by  
5 measuring the amount of subject protein present in a protein/antibody complex or, alternatively, by measuring the amount of remaining uncomplexed protein. The amount of protein may be detected by providing a labeled protein molecule.

[143] Immunoassays in the competitive binding format can be used for cross-reactivity determinations. For example, a protein of interest can be immobilized on a solid  
10 support. Proteins are added to the assay which compete with the binding of the antisera to the immobilized antigen. The ability of the above proteins to compete with the binding of the antisera to the immobilized protein is compared to that of the protein of interest. The percent cross-reactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% cross-reactivity with each of the proteins listed above are selected  
15 and pooled. The cross-reacting antibodies are optionally removed from the pooled antisera by immunoabsorption with the considered proteins, e.g., distantly related homologs.

[144] The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein, thought to be perhaps a protein of the present invention, to the immunogen protein. In order to make this  
20 comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required is less than 10 times the amount of the protein partially encoded by a sequence herein that is required, then the second protein is said to specifically bind to an antibody generated to an immunogen  
25 consisting of the target protein.

### 3. Other Assay Formats

[145] In a particularly preferred embodiment, western blot (immunoblot) analysis is used to detect and quantify the presence of a polypeptide of the invention in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis  
30 of molecular weight, transferring the separated proteins to a suitable solid support (such as, e.g., a nitrocellulose filter, a nylon filter, or a derivatized nylon filter) and incubating the sample with the antibodies that specifically bind the protein of interest. For example, the

antibodies specifically bind to a polypeptide of interest on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (e.g., labeled sheep anti-mouse antibodies) that specifically bind to the antibodies against the protein of interest.

- 5 [146] Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (e.g., antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (*see, Monroe et al. (1986) Amer. Clin. Prod. Rev. 5:34-41*).

#### 4. Labels

- 10 [147] The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well developed in the field of immunoassays and, in general, most labels useful in such methods can be applied  
15 to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (e.g., Dynabeads™), fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (e.g., <sup>3</sup>H, <sup>125</sup>I, <sup>35</sup>S, <sup>14</sup>C, or <sup>32</sup>P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others  
20 commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads.

- [148] The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on the sensitivity required, the ease of  
25 conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

- [149] Non-radioactive labels are often attached by indirect means. The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorescent compound. A variety of enzymes and fluorescent compounds can be  
30 used with the methods of the present invention and are well-known to those of skill in the art (for a review of various labeling or signal producing systems which may be used, *see, e.g., U.S. Patent No. 4,391,904*).

[150] Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and  
5 detecting the resulting fluorescence. The fluorescence may be detected visually, by means of photographic film, by the use of electronic detectors such as charge-coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple colorimetric labels may be detected directly by observing the color associated  
10 with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

[151] Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target  
15 antibodies. In this format, none of the components need to be labeled and the presence of the target antibody is detected by simple visual inspection.

## **VI. Screening for modulators of polypeptides and polynucleotides of the invention**

[152] Modulators of polypeptides or polynucleotides of the invention, i.e. agonists or antagonists of their activity or modulators of polypeptide or polynucleotide expression, are  
20 useful for treating a number of human diseases, including mood disorders or psychotic disorders. Administration of agonists, antagonists or other agents that modulate expression of the polynucleotides or polypeptides of the invention can be used to treat patients with mood disorders or psychotic disorders.

### **A. Screening methods**

[153] A number of different screening protocols can be utilized to identify agents  
25 that modulate the level of expression or activity of polypeptides and polynucleotides of the invention in cells, particularly mammalian cells, and especially human cells. In general terms, the screening methods involve screening a plurality of agents to identify an agent that modulates the polypeptide activity by binding to a polypeptide of the invention, modulating  
30 inhibitor binding to the polypeptide or activating expression of the polypeptide or polynucleotide, for example.

## 1. Binding Assays

[154] Preliminary screens can be conducted by screening for agents capable of binding to a polypeptide of the invention, as at least some of the agents so identified are likely modulators of polypeptide activity. The binding assays usually involve contacting a polypeptide of the invention with one or more test agents and allowing sufficient time for the protein and test agents to form a binding complex. Any binding complexes formed can be detected using any of a number of established analytical techniques. Protein binding assays include, but are not limited to, methods that measure co-precipitation, co-migration on non-denaturing SDS-polyacrylamide gels, and co-migration on Western blots (*see, e.g.*, Bennet and Yamamura, (1985) "Neurotransmitter, Hormone or Drug Receptor Binding Methods," in *Neurotransmitter Receptor Binding* (Yamamura, H. I., *et al.*, eds.), pp. 61-89. The protein utilized in such assays can be naturally expressed, cloned or synthesized.

[155] Binding assays are also useful, e.g., for identifying endogenous proteins that interact with a polypeptide of the invention. For example, antibodies, receptors or other molecules that bind a polypeptide of the invention can be identified in binding assays.

## 2. Expression Assays

[156] Certain screening methods involve screening for a compound that up or down-regulates the expression of a polypeptide or polynucleotide of the invention. Such methods generally involve conducting cell-based assays in which test compounds are contacted with one or more cells expressing a polypeptide or polynucleotide of the invention and then detecting an increase or decrease in expression (either transcript, translation product, or catalytic product). Some assays are performed with peripheral cells, or other cells, that express an endogenous polypeptide or polynucleotide of the invention.

[157] Polypeptide or polynucleotide expression can be detected in a number of different ways. As described *infra*, the expression level of a polynucleotide of the invention in a cell can be determined by probing the mRNA expressed in a cell with a probe that specifically hybridizes with a transcript (or complementary nucleic acid derived therefrom) of a polynucleotide of the invention. Probing can be conducted by lysing the cells and conducting Northern blots or without lysing the cells using in situ-hybridization techniques. Alternatively, a polypeptide of the invention can be detected using immunological methods in which a cell lysate is probed with antibodies that specifically bind to a polypeptide of the invention.



[158] Other cell-based assays are reporter assays conducted with cells that do not express a polypeptide or polynucleotide of the invention. Certain of these assays are conducted with a heterologous nucleic acid construct that includes a promoter of a polynucleotide of the invention that is operably linked to a reporter gene that encodes a detectable product. A number of different reporter genes can be utilized. Some reporters are inherently detectable. An example of such a reporter is green fluorescent protein that emits fluorescence that can be detected with a fluorescence detector. Other reporters generate a detectable product. Often such reporters are enzymes. Exemplary enzyme reporters include, but are not limited to,  $\beta$ -glucuronidase, chloramphenicol acetyl transferase (CAT); Alton and Vapnek (1979) *Nature* 282:864-869), luciferase,  $\beta$ -galactosidase, green fluorescent protein (GFP) and alkaline phosphatase (Toh, *et al.* (1980) *Eur. J. Biochem.* 182:231-238; and Hall *et al.* (1983) *J. Mol. Appl. Gen.* 2:101).

[159] In these assays, cells harboring the reporter construct are contacted with a test compound. A test compound that either activates the promoter by binding to it or triggers a cascade that produces a molecule that activates the promoter causes expression of the detectable reporter. Certain other reporter assays are conducted with cells that harbor a heterologous construct that includes a transcriptional control element that activates expression of a polynucleotide of the invention and a reporter operably linked thereto. Here, too, an agent that binds to the transcriptional control element to activate expression of the reporter or that triggers the formation of an agent that binds to the transcriptional control element to activate reporter expression, can be identified by the generation of signal associated with reporter expression.

[160] The level of expression or activity can be compared to a baseline value. As indicated above, the baseline value can be a value for a control sample or a statistical value that is representative of expression levels for a control population (e.g., healthy individuals not having or at risk for mood disorders or psychotic disorders). Expression levels can also be determined for cells that do not express a polynucleotide of the invention as a negative control. Such cells generally are otherwise substantially genetically the same as the test cells.

[161] A variety of different types of cells can be utilized in the reporter assays. Cells that express an endogenous polypeptide or polynucleotide of the invention include, e.g., brain cells, including cells from the cerebellum, anterior cingulate cortex, or dorsolateral prefrontal cortex. Such brain regions are part of brain circuits or pathways that are implicated in mood disorders. Cells that do not endogenously express polynucleotides of the invention

can be prokaryotic, but are preferably eukaryotic. The eukaryotic cells can be any of the cells typically utilized in generating cells that harbor recombinant nucleic acid constructs.

Exemplary eukaryotic cells include, but are not limited to, yeast, and various higher eukaryotic cells such as the COS, CHO and HeLa cell lines, and stem cells.

- 5 [162] Various controls can be conducted to ensure that an observed activity is authentic including running parallel reactions with cells that lack the reporter construct or by not contacting a cell harboring the reporter construct with test compound. Compounds can also be further validated as described below.

### 3. Catalytic activity

- 10 [163] Catalytic activity of polypeptides of the invention can be determined by measuring the production of enzymatic products or by measuring the consumption of substrates. Activity refers to either the rate of catalysis or the ability to the polypeptide to bind ( $K_m$ ) the substrate or release the catalytic product ( $K_d$ ).

- [164] Analysis of the activity of polypeptides of the invention are performed  
15 according to general biochemical analyses. Such assays include cell-based assays as well as *in vitro* assays involving purified or partially purified polypeptides or crude cell lysates. The assays generally involve providing a known quantity of substrate and quantifying product as a function of time.

### 4. Validation

- 20 [165] Agents that are initially identified by any of the foregoing screening methods can be further tested to validate the apparent activity. Preferably such studies are conducted with suitable animal models. The basic format of such methods involves administering a lead compound identified during an initial screen to an animal that serves as a model for humans and then determining if expression or activity of a polynucleotide or polypeptide of the  
25 invention is in fact upregulated. The animal models utilized in validation studies generally are mammals of any kind. Specific examples of suitable animals include, but are not limited to, primates, mice, and rats.

### 5. Animal models

- [166] Animal models of mental disorders also find use in screening for modulators.  
30 In one embodiment, rat models of depression (both chronic and acute), in which the rats are subjected to stress, are used for screening. In one embodiment, invertebrate models such as *Drosophila* models can be used, screening for modulators of *Drosophila* orthologs of the

human genes disclosed herein. In another embodiment, transgenic animal technology including gene knockout technology, for example as a result of homologous recombination with an appropriate gene targeting vector, or gene overexpression, will result in the absence, decreased or increased expression of a polynucleotide or polypeptide of the invention. The same technology can also be applied to make knockout cells. When desired, tissue-specific expression or knockout of a polynucleotide or polypeptide of the invention may be necessary. Transgenic animals generated by such methods find use as animal models of mental disorders and are useful in screening for modulators of mental disorders.

[167] Knockout cells and transgenic mice can be made by insertion of a marker gene or other heterologous gene into an endogenous gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting an endogenous polynucleotide of the invention with a mutated version of the polynucleotide, or by mutating an endogenous polynucleotide, e.g., by exposure to carcinogens.

[168] For development of appropriate stem cells, a DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic lesion (*see, e.g., Capecchi et al., Science* 244:1288 (1989)). Chimeric targeted mice can be derived according to Hogan *et al., Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory (1988) and *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed., IRL Press, Washington, D.C., (1987).

#### **B. Modulators of polypeptides or polynucleotides of the invention**

[169] The agents tested as modulators of the polypeptides or polynucleotides of the invention can be any small chemical compound, or a biological entity, such as a protein, sugar, nucleic acid or lipid. Alternatively, modulators can be genetically altered versions of a polypeptide or polynucleotide of the invention. Typically, test compounds will be small chemical molecules and peptides. Essentially any chemical compound can be used as a potential modulator or ligand in the assays of the invention, although most often compounds that can be dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to screen large chemical libraries by automating the assay steps and

providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs, Switzerland) and the like. Modulators also include agents designed to reduce the level of mRNA of the invention (e.g. antisense molecules, ribozymes, DNazymes and the like) or the level of translation from an mRNA.

[170] In one preferred embodiment, high throughput screening methods involve providing a combinatorial chemical or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds). Such “combinatorial chemical libraries” or “ligand libraries” are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional “lead compounds” or can themselves be used as potential or actual therapeutics.

[171] A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical “building blocks” such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

[172] Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.*, U.S. Patent 5,010,175, Furka, *Int. J. Pept. Prot. Res.* 37:487-493 (1991) and Houghton *et al.*, *Nature* 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (e.g., PCT Publication No. WO 91/19735), encoded peptides (e.g., PCT Publication WO 93/20242), random bio-oligomers (e.g., PCT Publication No. WO 92/00091), benzodiazepines (e.g., U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, *Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara *et al.*, *J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann *et al.*, *J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen

*et al.*, *J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho *et al.*, *Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al.*, *J. Org. Chem.* 59:658 (1994)), nucleic acid libraries (*see* Ausubel, Berger and Sambrook, all *supra*), peptide nucleic acid libraries (*see, e.g.*, U.S. Patent 5,539,083), antibody libraries (*see, e.g.*, Vaughn *et al.*, *Nature Biotechnology*, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (*see, e.g.*, Liang *et al.*, *Science*, 274:1520-1522 (1996) and U.S. Patent 5,593,853), small organic molecule libraries (*see, e.g.*, benzodiazepines, Baum C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like).

[173] Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY; Symphony, Rainin, Woburn, MA; 433A Applied Biosystems, Foster City, CA; 9050 Plus, Millipore, Bedford, MA). In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.*, ComGenex, Princeton, NJ; Tripos, Inc., St. Louis, MO; 3D Pharmaceuticals, Exton, PA; Martek Biosciences, Columbia, MD, etc.).

### C. Solid State and Soluble High Throughput Assays

[174] In the high throughput assays of the invention, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (e.g., 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100 to about 1500 different compounds. It is possible to assay several different plates per day; assay screens for up to about 6,000-20,000 different compounds are possible using the integrated systems of the invention. More recently, microfluidic approaches to reagent manipulation have been developed.

[175] The molecule of interest can be bound to the solid state component, directly or indirectly, via covalent or non-covalent linkage, e.g., via a tag. The tag can be any of a variety of components. In general, a molecule that binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest is attached to the solid support by interaction of the tag and the tag binder.

[176] A number of tags and tag binders can be used, based upon known molecular interactions well described in the literature. For example, where a tag has a natural binder, for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, etc.).

5 Antibodies to molecules with natural binders such as biotin are also widely available and appropriate tag binders (*see*, SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis MO).

[177] Similarly, any haptenic or antigenic compound can be used in combination with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies  
10 are commercially available and many additional antibodies are described in the literature. For example, in one common configuration, the tag is a first antibody and the tag binder is a second antibody which recognizes the first antibody. In addition to antibody-antigen interactions, receptor-ligand interactions are also appropriate as tag and tag-binder pairs, such as agonists and antagonists of cell membrane receptors (e.g., cell receptor-ligand interactions  
15 such as transferrin, c-kit, viral receptor ligands, cytokine receptors, chemokine receptors, interleukin receptors, immunoglobulin receptors and antibodies, the cadherin family, the integrin family, the selectin family, and the like; *see, e.g.*, Pigott & Power, *The Adhesion Molecule Facts Book I* (1993)). Similarly, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), intracellular receptors (e.g., which mediate the effects of various small  
20 ligands, including steroids, thyroid hormone, retinoids and vitamin D; peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with various cell receptors.

[178] Synthetic polymers, such as polyurethanes, polyesters, polycarbonates,  
25 polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this disclosure.

[179] Common linkers such as peptides, polyethers, and the like can also serve as  
30 tags, and include polypeptide sequences, such as poly-Gly sequences of between about 5 and 200 amino acids. Such flexible linkers are known to those of skill in the art. For example, poly(ethylene glycol) linkers are available from Shearwater Polymers, Inc., Huntsville,

Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

[180] Tag binders are fixed to solid substrates using any of a variety of methods currently available. Solid substrates are commonly derivatized or functionalized by exposing  
5 all or a portion of the substrate to a chemical reagent which fixes a chemical group to the surface which is reactive with a portion of the tag binder. For example, groups which are suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer  
10 arrays is well described in the literature (*see, e.g., Merrifield, J. Am. Chem. Soc.* 85:2149-2154 (1963) (describing solid phase synthesis of, e.g., peptides); Geysen *et al., J. Immun. Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank and Doring, *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences on cellulose disks); Fodor *et al., Science*, 251:767-777 (1991); Sheldon *et al., Clinical Chemistry* 39(4):718-719 (1993); and Kozal *et al., Nature Medicine* 2(7):753759  
15 (1996) (all describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

[181] The invention provides *in vitro* assays for identifying, in a high throughput  
20 format, compounds that can modulate the expression or activity of the polynucleotides or polypeptides of the invention. In a preferred embodiment, the methods of the invention include such a control reaction. For each of the assay formats described, "no modulator" control reactions that do not include a modulator provide a background level of binding activity.

[182] In some assays it will be desirable to have positive controls to ensure that the  
25 components of the assays are working properly. At least two types of positive controls are appropriate. First, a known activator of a polynucleotide or polypeptide of the invention can be incubated with one sample of the assay, and the resulting increase in signal resulting from an increased expression level or activity of polynucleotide or polypeptide determined  
30 according to the methods herein. Second, a known inhibitor of a polynucleotide or polypeptide of the invention can be added, and the resulting decrease in signal for the expression or activity can be similarly detected.

#### **D. Computer-Based Assays**

[183] Yet another assay for compounds that modulate the activity of a polypeptide or polynucleotide of the invention involves computer assisted drug design, in which a computer system is used to generate a three-dimensional structure of the polypeptide or polynucleotide based on the structural information encoded by its amino acid or nucleotide sequence. The input sequence interacts directly and actively with a pre-established algorithm in a computer program to yield secondary, tertiary, and quaternary structural models of the molecule. Similar analyses can be performed on potential receptors or binding partners of the polypeptides or polynucleotides of the invention. The models of the protein or nucleotide structure are then examined to identify regions of the structure that have the ability to bind, e.g., a polypeptide or polynucleotide of the invention. These regions are then used to identify polypeptides that bind to a polypeptide or polynucleotide of the invention.

[184] The three-dimensional structural model of a protein is generated by entering protein amino acid sequences of at least 10 amino acid residues or corresponding nucleic acid sequences encoding a potential receptor into the computer system. The amino acid sequences encoded by the nucleic acid sequences provided herein represent the primary sequences or subsequences of the proteins, which encode the structural information of the proteins. At least 10 residues of an amino acid sequence (or a nucleotide sequence encoding 10 amino acids) are entered into the computer system from computer keyboards, computer readable substrates that include, but are not limited to, electronic storage media (e.g., magnetic diskettes, tapes, cartridges, and chips), optical media (e.g., CD ROM), information distributed by internet sites, and by RAM. The three-dimensional structural model of the protein is then generated by the interaction of the amino acid sequence and the computer system, using software known to those of skill in the art.

[185] The amino acid sequence represents a primary structure that encodes the information necessary to form the secondary, tertiary, and quaternary structure of the protein of interest. The software looks at certain parameters encoded by the primary sequence to generate the structural model. These parameters are referred to as "energy terms," and primarily include electrostatic potentials, hydrophobic potentials, solvent accessible surfaces, and hydrogen bonding. Secondary energy terms include van der Waals potentials. Biological molecules form the structures that minimize the energy terms in a cumulative fashion. The computer program is therefore using these terms encoded by the primary structure or amino acid sequence to create the secondary structural model.



[186] The tertiary structure of the protein encoded by the secondary structure is then formed on the basis of the energy terms of the secondary structure. The user at this point can enter additional variables such as whether the protein is membrane bound or soluble, its location in the body, and its cellular location, e.g., cytoplasmic, surface, or nuclear. These variables along with the energy terms of the secondary structure are used to form the model of the tertiary structure. In modeling the tertiary structure, the computer program matches hydrophobic faces of secondary structure with like, and hydrophilic faces of secondary structure with like.

[187] Once the structure has been generated, potential ligand binding regions are identified by the computer system. Three-dimensional structures for potential ligands are generated by entering amino acid or nucleotide sequences or chemical formulas of compounds, as described above. The three-dimensional structure of the potential ligand is then compared to that of a polypeptide or polynucleotide of the invention to identify binding sites of the polypeptide or polynucleotide of the invention. Binding affinity between the protein and ligands is determined using energy terms to determine which ligands have an enhanced probability of binding to the protein.

[188] Computer systems are also used to screen for mutations, polymorphic variants, alleles and interspecies homologs of genes encoding a polypeptide or polynucleotide of the invention. Such mutations can be associated with disease states or genetic traits and can be used for diagnosis. As described above, GeneChip™ and related technology can also be used to screen for mutations, polymorphic variants, alleles and interspecies homologs. Once the variants are identified, diagnostic assays can be used to identify patients having such mutated genes. Identification of the mutated a polypeptide or polynucleotide of the invention involves receiving input of a first amino acid sequence of a polypeptide of the invention (or of a first nucleic acid sequence encoding a polypeptide of the invention), e.g., any amino acid sequence having at least 60%, optionally at least 70% or 85%, identity with the amino acid sequence of interest, or conservatively modified versions thereof. The sequence is entered into the computer system as described above. The first nucleic acid or amino acid sequence is then compared to a second nucleic acid or amino acid sequence that has substantial identity to the first sequence. The second sequence is entered into the computer system in the manner described above. Once the first and second sequences are compared, nucleotide or amino acid differences between the sequences are identified. Such sequences can represent allelic

differences in various polynucleotides, including SNPs and/or haplotypes, of the invention, and mutations associated with disease states and genetic traits.

## **VII. Compositions, Kits and Integrated Systems**

[189] The invention provides compositions, kits and integrated systems for practicing the assays described herein using polypeptides or polynucleotides of the invention, antibodies specific for polypeptides or polynucleotides of the invention, etc.

[190] The invention provides assay compositions for use in solid phase assays; such compositions can include, for example, one or more polynucleotides or polypeptides of the invention immobilized on a solid support, and a labeling reagent. In each case, the assay compositions can also include additional reagents that are desirable for hybridization. Modulators of expression or activity of polynucleotides or polypeptides of the invention can also be included in the assay compositions.

[191] The invention also provides kits for carrying out the therapeutic and diagnostic assays of the invention. The kits typically include a probe that comprises an antibody that specifically binds to polypeptides or polynucleotides of the invention, and a label for detecting the presence of the probe. The kits may include several polynucleotide sequences encoding polypeptides of the invention. Kits can include any of the compositions noted above, and optionally further include additional components such as instructions to practice a high-throughput method of assaying for an effect on expression of the genes encoding the polypeptides of the invention, or on activity of the polypeptides of the invention, one or more containers or compartments (e.g., to hold the probe, labels, or the like), a control modulator of the expression or activity of polypeptides of the invention, a robotic armature for mixing kit components or the like.

[192] The invention also provides integrated systems for high-throughput screening of potential modulators for an effect on the expression or activity of the polypeptides of the invention. The systems typically include a robotic armature which transfers fluid from a source to a destination, a controller which controls the robotic armature, a label detector, a data storage unit which records label detection, and an assay component such as a microtiter dish comprising a well having a reaction mixture or a substrate comprising a fixed nucleic acid or immobilization moiety.

[193] A number of robotic fluid transfer systems are available, or can easily be made from existing components. For example, a Zymate XP (Zymark Corporation; Hopkinton,

MA) automated robot using a Microlab 2200 (Hamilton; Reno, NV) pipetting station can be used to transfer parallel samples to 96 well microtiter plates to set up several parallel simultaneous STAT binding assays.

[194] Optical images viewed (and, optionally, recorded) by a camera or other recording device (e.g., a photodiode and data storage device) are optionally further processed in any of the embodiments herein, e.g., by digitizing the image and storing and analyzing the image on a computer. A variety of commercially available peripheral equipment and software is available for digitizing, storing and analyzing a digitized video or digitized optical image, e.g., using PC, MACINTOSH®, or UNIX® based (e.g., SUN® work station) computers.

[195] One conventional system carries light from the specimen field to a cooled charge-coupled device (CCD) camera, in common use in the art. A CCD camera includes an array of picture elements (pixels). The light from the specimen is imaged on the CCD. Particular pixels corresponding to regions of the specimen (e.g., individual hybridization sites on an array of biological polymers) are sampled to obtain light intensity readings for each position. Multiple pixels are processed in parallel to increase speed. The apparatus and methods of the invention are easily used for viewing any sample, e.g., by fluorescent or dark field microscopic techniques. Lasar based systems can also be used.

### VIII. Administration and Pharmaceutical compositions

[196] Modulators of the polynucleotides or polypeptides of the invention (e.g., antagonists or agonists) can be administered directly to a mammalian subject for modulation of activity of those molecules *in vivo*. Administration is by any of the routes normally used for introducing a modulator compound into ultimate contact with the tissue to be treated and is well known to those of skill in the art. Although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

[197] Diseases that can be treated include the following, which include the corresponding reference number from Morrison, *DSM-IV Made Easy*, 1995: Schizophrenia, Catatonic, Subchronic, (295.21); Schizophrenia, Catatonic, Chronic (295.22); Schizophrenia, Catatonic, Subchronic with Acute Exacerbation (295.23); Schizophrenia, Catatonic, Chronic with Acute Exacerbation (295.24); Schizophrenia, Catatonic, in Remission (295.55); Schizophrenia, Catatonic, Unspecified (295.20); Schizophrenia, Disorganized, Subchronic

- (295.11); Schizophrenia, Disorganized, Chronic (295.12); Schizophrenia, Disorganized, Subchronic with Acute Exacerbation (295.13); Schizophrenia, Disorganized, Chronic with Acute Exacerbation (295.14); Schizophrenia, Disorganized, in Remission (295.15); Schizophrenia, Disorganized, Unspecified (295.10); Schizophrenia, Paranoid, Subchronic (295.31); Schizophrenia, Paranoid, Chronic (295.32); Schizophrenia, Paranoid, Subchronic with Acute Exacerbation (295.33); Schizophrenia, Paranoid, Chronic with Acute Exacerbation (295.34); Schizophrenia, Paranoid, in Remission (295.35); Schizophrenia, Paranoid, Unspecified (295.30); Schizophrenia, Undifferentiated, Subchronic (295.91); Schizophrenia, Undifferentiated, Chronic (295.92); Schizophrenia, Undifferentiated, Subchronic with Acute Exacerbation (295.93); Schizophrenia, Undifferentiated, Chronic with Acute Exacerbation (295.94); Schizophrenia, Undifferentiated, in Remission (295.95); Schizophrenia, Undifferentiated, Unspecified (295.90); Schizophrenia, Residual, Subchronic (295.61); Schizophrenia, Residual, Chronic (295.62); Schizophrenia, Residual, Subchronic with Acute Exacerbation (295.63); Schizophrenia, Residual, Chronic with Acute Exacerbation (295.94); Schizophrenia, Residual, in Remission (295.65); Schizophrenia, Residual, Unspecified (295.60); Delusional (Paranoid) Disorder (297.10); Brief Reactive Psychosis (298.80); Schizophreniform Disorder (295.40); Schizoaffective Disorder (295.70); Induced Psychotic Disorder (297.30); Psychotic Disorder NOS (Atypical Psychosis) (298.90); Personality Disorders, Paranoid (301.00); Personality Disorders, Schizoid (301.20); Personality Disorders, Schizotypal (301.22); Personality Disorders, Antisocial (301.70); Personality Disorders, Borderline (301.83) and bipolar disorders, maniac, hypomaniac, dysthymic or cyclothymic disorders, substance-induced mood disorders, major depression, psychotic disorders, including paranoid psychosis, catatonic psychosis, delusional psychosis, having schizoaffective disorder, and substance-induced psychotic disorder.
- 25 **[198]** In some embodiments, modulators of polynucleotides or polypeptides of the invention can be combined with other drugs useful for treating mental disorders including useful for treating mood disorders, e.g., schizophrenia, bipolar disorders, or major depression. In some preferred embodiments, pharmaceutical compositions of the invention comprise a modulator of a polypeptide of polynucleotide of the invention combined with at least one of
- 30 the compounds useful for treating schizophrenia, bipolar disorder, or major depression, e.g., such as those described in U.S. Patent Nos. 6,297,262; 6,284,760; 6,284,771; 6,232,326; 6,187,752; 6,117,890; 6,239,162 or 6,166,008.

[199] The pharmaceutical compositions of the invention may comprise a pharmaceutically acceptable carrier. Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable  
5 formulations of pharmaceutical compositions of the present invention (*see, e.g., Remington's Pharmaceutical Sciences*, 17<sup>th</sup> ed. 1985)).

[200] The modulators (e.g., agonists or antagonists) of the expression or activity of the a polypeptide or polynucleotide of the invention, alone or in combination with other suitable components, can be made into aerosol formulations (i.e., they can be "nebulized") to  
10 be administered via inhalation or in compositions useful for injection. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

[201] Formulations suitable for administration include aqueous and non-aqueous solutions, isotonic sterile solutions, which can contain antioxidants, buffers, bacteriostats, and  
15 solutes that render the formulation isotonic, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, orally, nasally, topically, intravenously, intraperitoneally, or intrathecally. The formulations of compounds can be presented in unit-dose or multi-dose sealed containers,  
20 such as ampoules and vials. Solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. The modulators can also be administered as part of a prepared food or drug.

[202] The dose administered to a patient, in the context of the present invention should be sufficient to effect a beneficial response in the subject over time. The optimal dose  
25 level for any patient will depend on a variety of factors including the efficacy of the specific modulator employed, the age, body weight, physical activity, and diet of the patient, on a possible combination with other drugs, and on the severity of the mental disorder. The size of the dose also will be determined by the existence, nature, and extent of any adverse side effects that accompany the administration of a particular compound or vector in a particular  
30 subject.

[203] In determining the effective amount of the modulator to be administered a physician may evaluate circulating plasma levels of the modulator, modulator toxicity, and

the production of anti-modulator antibodies. In general, the dose equivalent of a modulator is from about 1 ng/kg to 10 mg/kg for a typical subject.

[204] For administration, modulators of the present invention can be administered at a rate determined by the LD-50 of the modulator, and the side effects of the modulator at various concentrations, as applied to the mass and overall health of the subject. Administration can be accomplished via single or divided doses.

## IX. Gene Therapy Applications

[205] A variety of human diseases can be treated by therapeutic approaches that involve stably introducing a gene into a human cell such that the gene is transcribed and the gene product is produced in the cell. Diseases amenable to treatment by this approach include inherited diseases, including those in which the defect is in a single or multiple genes. Gene therapy is also useful for treatment of acquired diseases and other conditions. For discussions on the application of gene therapy towards the treatment of genetic as well as acquired diseases, *see*, Miller, *Nature* 357:455-460 (1992); and Mulligan, *Science* 260:926-932 (1993).

[206] In the context of the present invention, gene therapy can be used for treating a variety of disorders and/or diseases in which the polynucleotides and polypeptides of the invention has been implicated. For example, compounds, including polynucleotides, can be identified by the methods of the present invention as effective in treating a mental disorder. Introduction by gene therapy of these polynucleotides can then be used to treat, e.g., mental disorders including mood disorders and psychotic disorders.

### A. Vectors for Gene Delivery

[207] For delivery to a cell or organism, the polynucleotides of the invention can be incorporated into a vector. Examples of vectors used for such purposes include expression plasmids capable of directing the expression of the nucleic acids in the target cell. In other instances, the vector is a viral vector system wherein the nucleic acids are incorporated into a viral genome that is capable of transfecting the target cell. In a preferred embodiment, the polynucleotides can be operably linked to expression and control sequences that can direct expression of the gene in the desired target host cells. Thus, one can achieve expression of the nucleic acid under appropriate conditions in the target cell.

## **B. Gene Delivery Systems**

[208] Viral vector systems useful in the expression of the nucleic acids include, for example, naturally occurring or recombinant viral vector systems. Depending upon the particular application, suitable viral vectors include replication competent, replication  
5 deficient, and conditionally replicating viral vectors. For example, viral vectors can be derived from the genome of human or bovine adenoviruses, vaccinia virus, herpes virus, adeno-associated virus, minute virus of mice (MVM), HIV, sindbis virus, and retroviruses (including but not limited to Rous sarcoma virus), and MoMLV. Typically, the genes of interest are inserted into such vectors to allow packaging of the gene construct, typically with  
10 accompanying viral DNA, followed by infection of a sensitive host cell and expression of the gene of interest.

[209] As used herein, "gene delivery system" refers to any means for the delivery of a nucleic acid of the invention to a target cell. In some embodiments of the invention, nucleic acids are conjugated to a cell receptor ligand for facilitated uptake (e.g., invagination of  
15 coated pits and internalization of the endosome) through an appropriate linking moiety, such as a DNA linking moiety (Wu *et al.*, *J. Biol. Chem.* 263:14621-14624 (1988); WO 92/06180). For example, nucleic acids can be linked through a polylysine moiety to asialo-oromucoid, which is a ligand for the asialoglycoprotein receptor of hepatocytes.

[210] Similarly, viral envelopes used for packaging gene constructs that include the  
20 nucleic acids of the invention can be modified by the addition of receptor ligands or antibodies specific for a receptor to permit receptor-mediated endocytosis into specific cells (*see, e.g.*, WO 93/20221, WO 93/14188, and WO 94/06923). In some embodiments of the invention, the DNA constructs of the invention are linked to viral proteins, such as adenovirus particles, to facilitate endocytosis (Curiel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*  
25 88:8850-8854 (1991)). In other embodiments, molecular conjugates of the instant invention can include microtubule inhibitors (WO/9406922), synthetic peptides mimicking influenza virus hemagglutinin (Plank *et al.*, *J. Biol. Chem.* 269:12918-12924 (1994)), and nuclear localization signals such as SV40 T antigen (WO93/19768).

[211] Retroviral vectors are also useful for introducing the nucleic acids of the  
30 invention into target cells or organisms. Retroviral vectors are produced by genetically manipulating retroviruses. The viral genome of retroviruses is RNA. Upon infection, this genomic RNA is reverse transcribed into a DNA copy which is integrated into the chromosomal DNA of transduced cells with a high degree of stability and efficiency. The

integrated DNA copy is referred to as a provirus and is inherited by daughter cells as is any other gene. The wild type retroviral genome and the proviral DNA have three genes: the *gag*, the *pol* and the *env* genes, which are flanked by two long terminal repeat (LTR) sequences. The *gag* gene encodes the internal structural (nucleocapsid) proteins; the *pol* gene encodes the RNA directed DNA polymerase (reverse transcriptase); and the *env* gene encodes viral envelope glycoproteins. The 5' and 3' LTRs serve to promote transcription and polyadenylation of virion RNAs. Adjacent to the 5' LTR are sequences necessary for reverse transcription of the genome (the tRNA primer binding site) and for efficient encapsulation of viral RNA into particles (the Psi site) (see, Mulligan, In: *Experimental Manipulation of Gene Expression*, Inouye (ed), 155-173 (1983); Mann *et al.*, *Cell* 33:153-159 (1983); Cone and Mulligan, *Proceedings of the National Academy of Sciences, U.S.A.*, 81:6349-6353 (1984)).

[212] The design of retroviral vectors is well known to those of ordinary skill in the art. In brief, if the sequences necessary for encapsidation (or packaging of retroviral RNA into infectious virions) are missing from the viral genome, the result is a *cis*-acting defect which prevents encapsidation of genomic RNA. However, the resulting mutant is still capable of directing the synthesis of all virion proteins. Retroviral genomes from which these sequences have been deleted, as well as cell lines containing the mutant genome stably integrated into the chromosome are well known in the art and are used to construct retroviral vectors. Preparation of retroviral vectors and their uses are described in many publications including, e.g., European Patent Application EPA 0 178 220; U.S. Patent 4,405,712, Gilboa *Biotechniques* 4:504-512 (1986); Mann *et al.*, *Cell* 33:153-159 (1983); Cone and Mulligan *Proc. Natl. Acad. Sci. USA* 81:6349-6353 (1984); Eglitis *et al.* *Biotechniques* 6:608-614 (1988); Miller *et al.* *Biotechniques* 7:981-990 (1989); Miller (1992) *supra*; Mulligan (1993), *supra*; and WO 92/07943.

[213] The retroviral vector particles are prepared by recombinantly inserting the desired nucleotide sequence into a retrovirus vector and packaging the vector with retroviral capsid proteins by use of a packaging cell line. The resultant retroviral vector particle is incapable of replication in the host cell but is capable of integrating into the host cell genome as a proviral sequence containing the desired nucleotide sequence. As a result, the patient is capable of producing, for example, a polypeptide or polynucleotide of the invention and thus restore the cells to a normal phenotype.

[214] Packaging cell lines that are used to prepare the retroviral vector particles are typically recombinant mammalian tissue culture cell lines that produce the necessary viral



structural proteins required for packaging, but which are incapable of producing infectious virions. The defective retroviral vectors that are used, on the other hand, lack these structural genes but encode the remaining proteins necessary for packaging. To prepare a packaging cell line, one can construct an infectious clone of a desired retrovirus in which the packaging site has been deleted. Cells comprising this construct will express all structural viral proteins, but the introduced DNA will be incapable of being packaged. Alternatively, packaging cell lines can be produced by transforming a cell line with one or more expression plasmids encoding the appropriate core and envelope proteins. In these cells, the *gag*, *pol*, and *env* genes can be derived from the same or different retroviruses.

[215] A number of packaging cell lines suitable for the present invention are also available in the prior art. Examples of these cell lines include Crip, GPE86, PA317 and PG13 (*see Miller et al., J. Virol.* 65:2220-2224 (1991)). Examples of other packaging cell lines are described in Cone and Mulligan *Proceedings of the National Academy of Sciences, USA*, 81:6349-6353 (1984); Danos and Mulligan *Proceedings of the National Academy of Sciences, USA*, 85:6460-6464 (1988); Eglitis *et al.* (1988), *supra*; and Miller (1990), *supra*.

[216] Packaging cell lines capable of producing retroviral vector particles with chimeric envelope proteins may be used. Alternatively, amphotropic or xenotropic envelope proteins, such as those produced by PA317 and GPX packaging cell lines may be used to package the retroviral vectors.

[217] In some embodiments of the invention, an antisense polynucleotide is administered which hybridizes to a gene encoding a polypeptide of the invention. The antisense polypeptide can be provided as an antisense oligonucleotide (*see, e.g., Murayama et al., Antisense Nucleic Acid Drug Dev.* 7:109-114 (1997)). Genes encoding an antisense nucleic acid can also be provided; such genes can be introduced into cells by methods known to those of skill in the art. For example, one can introduce an antisense nucleotide sequence in a viral vector, such as, for example, in hepatitis B virus (*see, e.g., Ji et al., J. Viral Hepat.* 4:167-173 (1997)), in adeno-associated virus (*see, e.g., Xiao et al., Brain Res.* 756:76-83 (1997)), or in other systems including, but not limited, to an HVJ (Sendai virus)-liposome gene delivery system (*see, e.g., Kaneda et al., Ann. NY Acad. Sci.* 811:299-308 (1997)), a "peptide vector" (*see, e.g., Vidal et al., CR Acad. Sci III* 32:279-287 (1997)), as a gene in an episomal or plasmid vector (*see, e.g., Cooper et al., Proc. Natl. Acad. Sci. U.S.A.* 94:6450-6455 (1997), Yew *et al. Hum Gene Ther.* 8:575-584 (1997)), as a gene in a peptide-DNA aggregate (*see, e.g., Niidome et al., J. Biol. Chem.* 272:15307-15312 (1997)), as "naked

DNA" (see, e.g., U.S. patent Nos. 5,580,859 and 5,589,466), in lipidic vector systems (see, e.g., Lee *et al.*, *Crit Rev Ther Drug Carrier Syst.* 14:173-206 (1997)), polymer coated liposomes (U.S. patent Nos. 5,213,804 and 5,013,556), cationic liposomes (Epanand *et al.*, U.S. patent Nos. 5,283,185; 5,578,475; 5,279,833; and 5,334,761), gas filled microspheres (U.S. patent No. 5,542,935), ligand-targeted encapsulated macromolecules (U.S. patent Nos. 5,108,921; 5,521,291; 5,554,386; and 5,166,320). In another embodiment, conditional expression systems, such as those typified by the tet-regulated systems and the RU-486 system, can be used (see, e.g., Gossen & Bujard, *PNAS* 89:5547 (1992); Oligino *et al.*, *Gene Ther.* 5:491-496 (1998); Wang *et al.*, *Gene Ther.* 4:432-441 (1997); Neering *et al.*, *Blood* 88:1147-1155 (1996); and Rendahl *et al.*, *Nat. Biotechnol.* 16:757-761 (1998)). These systems impart small molecule control on the expression of the target gene(s) of interest.

### C. Pharmaceutical Formulations

[218] When used for pharmaceutical purposes, the vectors used for gene therapy are formulated in a suitable buffer, which can be any pharmaceutically acceptable buffer, such as phosphate buffered saline or sodium phosphate/sodium sulfate, Tris buffer, glycine buffer, sterile water, and other buffers known to the ordinarily skilled artisan such as those described by Good *et al.* *Biochemistry* 5:467 (1966).

[219] The compositions can additionally include a stabilizer, enhancer, or other pharmaceutically acceptable carriers or vehicles. A pharmaceutically acceptable carrier can contain a physiologically acceptable compound that acts, for example, to stabilize the nucleic acids of the invention and any associated vector. A physiologically acceptable compound can include, for example, carbohydrates, such as glucose, sucrose or dextrans; antioxidants, such as ascorbic acid or glutathione; chelating agents; low molecular weight proteins or other stabilizers or excipients. Other physiologically acceptable compounds include wetting agents, emulsifying agents, dispersing agents, or preservatives, which are particularly useful for preventing the growth or action of microorganisms. Various preservatives are well known and include, for example, phenol and ascorbic acid. Examples of carriers, stabilizers, or adjuvants can be found in Remington's *Pharmaceutical Sciences*, Mack Publishing Company, Philadelphia, PA, 17th ed. (1985).

### D. Administration of Formulations

[220] The formulations of the invention can be delivered to any tissue or organ using any delivery method known to the ordinarily skilled artisan. In some embodiments of the

invention, the nucleic acids of the invention are formulated in mucosal, topical, and/or buccal formulations, particularly mucoadhesive gel and topical gel formulations. Exemplary permeation enhancing compositions, polymer matrices, and mucoadhesive gel preparations for transdermal delivery are disclosed in U.S. Patent No. 5,346,701.

## 5 E. Methods of Treatment

[221] The gene therapy formulations of the invention are typically administered to a cell. The cell can be provided as part of a tissue, such as an epithelial membrane, or as an isolated cell, such as in tissue culture. The cell can be provided *in vivo*, *ex vivo*, or *in vitro*.

10 [222] The formulations can be introduced into the tissue of interest *in vivo* or *ex vivo* by a variety of methods. In some embodiments of the invention, the nucleic acids of the invention are introduced into cells by such methods as microinjection, calcium phosphate precipitation, liposome fusion, or biolistics. In further embodiments, the nucleic acids are taken up directly by the tissue of interest.

15 [223] In some embodiments of the invention, the nucleic acids of the invention are administered *ex vivo* to cells or tissues explanted from a patient, then returned to the patient. Examples of *ex vivo* administration of therapeutic gene constructs include Nolta *et al.*, *Proc Natl. Acad. Sci. USA* 93(6):2414-9 (1996); Koc *et al.*, *Seminars in Oncology* 23 (1):46-65 (1996); Raper *et al.*, *Annals of Surgery* 223(2):116-26 (1996); Dalesandro *et al.*, *J. Thorac. Cardi. Surg.*, 11(2):416-22 (1996); and Makarov *et al.*, *Proc. Natl. Acad. Sci. USA* 93(1):402-20 6 (1996).

## X. Diagnosis of mood disorders and psychotic disorders

[224] The present invention also provides methods of diagnosing mood disorders (such as major depression or bipolar disorder), psychotic disorders (such as schizophrenia) Diagnosis involves determining the level of a polypeptide or polynucleotide of the invention 25 in a patient and then comparing the level to a baseline or range. Typically, the baseline value is representative of a polypeptide or polynucleotide of the invention in a healthy person not suffering from a mood disorder or psychotic disorder or under the effects of medication or other drugs. Variation of levels of a polypeptide or polynucleotide of the invention from the baseline range (either up or down) indicates that the patient has a mood disorder or psychotic 30 disorder or at risk of developing at least some aspects of a mood disorder or psychotic disorder. In some embodiments, the level of a polypeptide or polynucleotide of the invention are measured by taking a blood, urine or tissue sample from a patient and measuring the

amount of a polypeptide or polynucleotide of the invention in the sample using any number of detection methods, such as those discussed herein, e.g., SNPs or haplotypes associated with this genes.

[225] In some embodiments, the level of the enzymatic product of a polypeptide or polynucleotide of the invention is measured and compared to a baseline value of a healthy person or persons. Modulated levels of the product compared to the baseline indicates that the patient has a mood disorder or psychotic disorder or is at risk of developing at least some aspects of a mood disorder or psychotic disorder. Patient samples, for example, can be blood, saliva, CSF, urine or tissue samples.

[226] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

## EXAMPLES

### **Example 1: Identification of genes dysregulated in mood disorders**

[227] A total of twenty mood disorder brains (9 bipolar and 11 major depression disorder patients) with twenty control brains were used in this study. Each brain pair (case and control) was matched on the basis of gender, age, and postmortem interval. Three brain regions, dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex (AnCg) and the cerebellum (CB) were extracted for RNA and subjected to microarray analysis using Affymetrix oligonucleotide GeneChips™. Each RNA sample was subjected to two independent analyses. The results were analyzed using multiple statistical tools and algorithms with various stringencies. Real time PCR analysis was used to confirm differential gene expression for selected genes. The genes identified using this study are listed in Tables 1, 2, and 3. Furthermore, biochemical pathways associated with the differentially expressed genes were identified (see Figures 1-5).

[228] The two cortical regions DLPFC and AnCg had similar gene expression profiles in controls but differed significantly in MDD and BP, demonstrating distinct gene expression profiles. BP subject showed more changes in AnCg compared to DLPFC whereas MDD show less profound changes in both cortical regions but had greater effects in the

DLPFC than in the AnCg. For BP, several candidate genes were located in chromosomal region 15q11-13, which is associated with the Prader-Willi syndrome (see Figures 6-8).

### **Example 2: Identification of additional genes dysregulated in mood disorders**

[229] The RNA from three brain regions, dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex (AnCg) and the cerebellum (CB) from deceased patients diagnosed with bipolar disease or major depression, and matched controls were extracted and subjected to microarray analysis using Affymetrix oligonucleotide GeneChips™. The patient's particular conditions in their terminal phase (agonal factors, e.g., seizure, coma, hypoxia, dehydration, and pyrexia) and the conditions of the brain tissue after death (postmortem factors, e.g., postmortem interval, and freezer interval) are two major influences on RNA preservation in postmortem brain tissue. Brain pH has been evaluated as an indicator for agonal status, and as an indicator of RNA preservation. The effect of agonal factors and pH were taken into account for quality control of the RNA. Two RNA samples were subjected to independent analyses. The results were analyzed using multiple statistical tools and algorithms with various stringencies. The 967 genes identified using this study are listed in Table 4. Real time PCR analysis was used to confirm differential gene expression for selected genes. Real time PCR confirmation of differential gene expression for selected genes is listed in Table 5.

[230] Furthermore, biochemical pathways associated with the differentially expressed genes were identified. In particular, cortical areas in BP patients showed activation of several pathways, including the proteasome pathway, the oxidative phosphorylation pathway, the ATP synthesis pathway, and chaperones (i.e., heat shock proteins). In addition, signaling pathways dysregulated in BP include, e.g., G-coupled protein receptors, the phosphatidylinositol pathway, the cAMP pathway, the mitogen activated protein kinase pathway, cytoskeletal systems, and the cortical GABA and glutamate systems. In MD, dysregulated genes includes genes involved in transmission of nerve impulses, neurogenesis, and the fibroblast growth factor system (FGF). (see Figures 10-12). Gene ontology (i.e., genetic signatures) for BP and MD can conveniently be used in developing diagnostic and therapeutic regimens for mood disorders.

**Example 3: Identification of additional genes dysregulated in mood disorders using rat models of depression and anti-depressant treatment**

[231] Rats were exposed to chronic unpredictable stress treatments in parallel with chronic anti-depressants treatment (e.g., the tricyclic antidepressant desipramine and the specific serotonin reuptake inhibitor fluoxetine). Saline treated stressed rats (SS) and saline treated non-stressed rats (SN) were used as controls. In particular, saline treated stressed rats (SS) were compared to desipramine treated stressed rats (DS); saline treated stressed rats (SS) were compared to fluoxetine treated stressed rats (FS); saline treated non-stressed rats (SN) were compared to desipramine treated non-stressed rats (DN); saline treated non-stressed rats (SN) were compared to fluoxetine treated non-stressed rats (FN); and saline treated stressed rats (SS) were compared to saline treated non-stressed rats (SN). Gene expression changes in rat cortex following treatment were measured. The genes identified in this study are shown in Table 6. This data suggests that different classes of antidepressants, i.e., antidepressants with apparently different mechanisms of action may act through a common biochemical pathway.

[232] The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, databases, Genbank sequences, GO terms, patents, and patent applications cited herein are hereby incorporated by reference.